Epigenetic Regulators as New Therapeutic Targets in Pediatric Sarcoma

---

Dissertation

zur

Erlangung der naturwissenschaftlichen Doktorwürde
(Dr. sc. nat.)

vorgelegt der

Mathematisch-naturwissenschaftlichen Fakultät

der

Universität Zürich

von

Joana Graça Marques

von/aus

Portugal

Promotionskommission

Prof. Dr. Beat W. Schäfer (Vorsitz und Leitung der Dissertation)
Prof. Dr. Amedeo Caflisch
Dr. Raffaella Santoro
Dr. César Fernández

Zürich, 2018

Para o meu ídolo,

O meu herói,

O meu avô,

José Maria

Table of Contents

Abstract

Tumorigenesis is driven by changes in the expression profile of healthy cells which are caused mainly by the disruption of signaling pathways or by oncogenic transcription factors. Then, cancer cells become addicted to high levels of aberrant transcription and to the expression of certain oncogenes.

Pediatric malignancies are characterized by a low mutational burden and are commonly caused by chromosomal translocations that originate fusion transcription factors (TFs). Hence, they are good candidates to study aberrant transcription. Fusion-positive rhabdomyosarcoma (FP-RMS) and Ewing sarcoma (ES) are two examples of this. Both tumors are driven by fusion TFs, PAX3-FOXO1 and EWS-FLI1 respectively, and are highly dependent on their transcriptional signature.

One way of interfering with transcription driven by oncogenic TFs is by targeting their epigenetic cofactors. Epigenetics studies alterations in gene expression that do not arise from changes in the DNA sequence, but rather from post-translational modifications on histones and DNA methylation which signal to effector enzymes that modulate gene expression. Thus, targeting these epigenetic enzymatic motors is a good strategy to interfere with oncogenic transcription and for cancer therapy. For example, the inhibition of BET-bromodomains deeply affects the transcriptional signatures of many oncogenes and impacts several hematological malignancies and solid tumors.

Here, we studied chromatin remodeling as a means to impair oncogenic transcription and cancer development. Chromatin remodelers use energy from ATP hydrolysis to move nucleosomes along the DNA and hence determine DNA accessibility to the transcription machinery. Despite their prominent role in transcription regulation, they have been so far neglected as therapeutic targets. In this study, we focused on the chromatin remodeler CHD4 and on its effect in FP-RMS and ES to highlight the role of chromatin remodeling in cancer.

We started by studying the epigenetic landscape of FP-RMS. We showed that the driver oncogene, PAX3-FOXO1, establishes and binds to enhancers and super-enhancers (SEs) from where it regulates gene expression together with the epigenetic cofactors BRD4 and CHD4. We showed that CHD4 presence in SEs is crucial to keep these regulatory elements open and permissive to the binding of the fusion protein and BRD4. Consequently, silencing of CHD4 deeply impaired the oncogenic program of PAX3-FOXO1, caused tumor cell death in vitro and tumor regression in vivo. Furthermore, we described for the first time an association between CHD4 and BRD4, which suggests an interdependency between these two epigenetic factors in the regulation of SE-driven gene expression in FP-RMS.

Ewing sarcoma, as previously stated, has similar characteristics to FP-RMS. Its driver, EWS-FLI1, also establishes de novo enhancers from where it controls aberrant gene expression. In parallel to FP-RMS, ES is highly dependent on BRD4 and studies have suggested that CHD4 interacts with EWS-FLI1 to influence its repressive action. Here, we complemented this work by characterizing a new function of CHD4 in ES cell proliferation and tumor growth in vivo.

Besides FP-RMS and ES, many tumors, like breast cancer, colorectal cancer, and acute myeloid leukemia, are dependent on CHD4. Interestingly, we and others observed that healthy cells are usually not affected by CHD4 silencing, which makes CHD4 a very interesting new cancer therapeutic target. In addition, we analyzed two genome-wide cancer dependency databases and observed that, amongst all SNF2-like ATPases, CHD4 depletion affected the largest number of tumors, similarly to what has been described for BRD4.

Motivated by these results, we decided to start developing the first-in-class CHD4 small molecule inhibitor. To do so, we focused on the reader domains of CHD4, named plant homeodomains (PHDs), and gathered a library of candidate small molecule inhibitors through *in silico* structure-based drug design. After biophysical validation of the candidate compounds by NMR experiments, we were able to identify a small molecule compound that could interact with the PHD2 domain of CHD4, although weakly. Thus, further optimization will be required to increase the binding affinity of this compound.

In conclusion, we reveal here a new role of CHD4 as a regulator of SE-driven transcription. We show that this remodeler conditions chromatin accessibility to aberrant TFs and their cofactors, hence influencing oncogenic gene expression and tumor development. We also describe a co-dependency between CHD4 and BRD4, which might explain their vast relevance in oncogenic transcription regulation. In addition, we show a broad, and possibly specific, tumor dependency to CHD4 and reveal its promising features as a new tumor target. With this work, we hope to highlight the importance of chromatin remodelers in tumor development and to place them in the epigenetic drug development pipeline.

Chapter One:
Introduction and Aim

Cancer

The word cancer comes from the Greek *karkinos* (crab). Hippocrates named it like that since the finger-like spreading projections seen in cancers resemble the shape of a crab[1,2]. Today, the Oxford Dictionary describes the disease as "an uncontrolled division of abnormal cells in a part of the body"[3].

Despite being an old disease, cancer is still the second cause of death and morbidity in Europe, after cardiovascular diseases[4,5,6]. The oldest description of cancer, the Edwin Smith Papyrus, comes from Egypt and dates to 3000 BC. There, 8 cases of breast cancer are described and the disease is identified as non-treatable[1].

In 2012, roughly 3.4 million new cases were diagnosed and 8.2 million people died of this disease[4]. In the USA, half of the men and one third of the women are estimated to develop cancer throughout their lifetime[2]. Europe has a particular challenging situation since, although it accounts for only one-ninth of the world population, it contains a quarter of the total number cancer cases[7].

1. Tumorigenesis

Cancers can be divided into inherited or sporadic[8]. The inherited ones account for 5% of all tumors[8] and are originated by germline mutations[9]. The well-known BRC1/2 mutations and the hereditary breast or ovarian cancer illustrate an example inherited cancer[8,10]. In sporadic cancers, genetic alterations have different origins such as viral or environmental[8]. Nevertheless, both inherited and sporadic cancers go through a multistep process in which cells suffer sequential genetic alterations, in oncogenes and/or tumor suppressor genes, that confer growth advantages which then drive the consecutive clonal expansion of malignant cells. This process is called tumorigenesis and ultimately leads to cell transformation[11].

Oncogenes are genes that promote cell proliferation, like growth factor receptors, while tumor suppressor genes restrict cell growth, like pro-apoptotic proteins[9]. The hallmarks of cancer, depicted in Fig.1, describe the biologic capabilities acquired by the cancer cell that result from the imbalance between the activity of oncogenes and tumor suppressor genes[12].

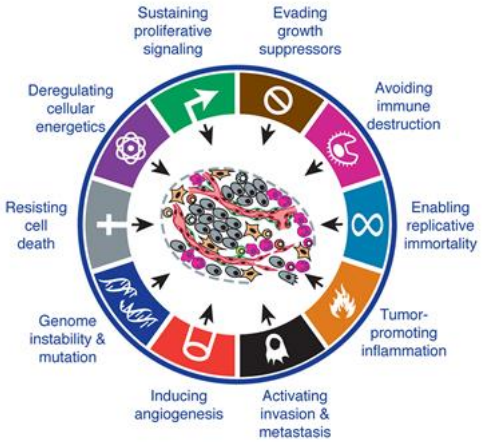

Fig. 1 Established and emerging hallmarks of cancer[12].

Not all mutations acquired are equally important. Gatekeeping mutations initiate the tumorigenesis process and allow a first clonal expansion. Other mutations that confer selective growth advantages are denominated driver mutations, while mutations that do not influence the neoplastic process are called passenger[13].

The clonality of cancer, meaning if tumors arise from genetic alterations in a single cell or in two or more cells, is still a subject of debate[14]. Nonetheless, cancer is an extremely heterogeneous disease, presenting intratumoral, intrametastic and intermetastatic diversity, which makes therapy very challenging[13].

2. The cancer genome

The cancer genome is highly aberrant. A normal diploid cell has a total of 46 chromosomes, but neuroblastoma cells can have up to 80. Many point mutations, viral insertions, chromosomal or gene amplifications, deletions or rearrangements are found in the genome of cancer cells. Point mutations can activate oncogenes or inactivate tumor suppressor genes, and gene amplifications allow the tumor to overexpress oncogenes[9]. These genomic aberrations allow cancer cells to evade cell death and maintain a high level of proliferation. Cancer cells can survive under extreme stress conditions unlike healthy cells. Mutations in KRAS and BRAF allow cancer cells to grow in low glucose concentrations. Mutations in EGFR, RAS, PIK3CA, and PTEN keep growth factor related pathways active in the absence of growth factors. Deregulation of cell cycle and apoptosis, through mutations in MYC or BCL2, also confer growth advantages[13].

Understanding how mutations in the genome alter normal cellular processes is crucial to further guide cancer therapy.

3. Cancer therapy

The three pillars of cancer therapy are chemotherapy, radiotherapy, and surgery.

Chemotherapy was introduced in the 1930s, and most of the drugs in use today were developed in the 1970s. This treatment modality consists in the combination of drugs, that act in different stages of the cell cycle, to more effectively kill the cancer cells. Unfortunately, the lack of specificity of these drugs causes severe side effects[15].

In the 1960s surgery and radiotherapy, were the basis of solid tumor treatment. Surgery has evolved from radical removal of tumor and adjacent tissue to less extensive and invasive techniques, like laparoscopy. Radiotherapy has also evolved, and now smaller and more localized radiation dosages are used. The latest advance is the usage of radiolabeled molecules, which fight cancer by delivering targeted radiation to specific receptor-bearing cells[15].

The advances in chemotherapeutical regiments, surgery techniques and radiotherapy have considerably enhanced patient's outcome (Fig.2)[16]. Nonetheless, we seem to have reached a plateau in the improvement of cancer survival. Hence, more specific and efficacious therapies are highly demanded. Targeted therapy, immunotherapy, and gene therapy hold great promise in cancer treatment.

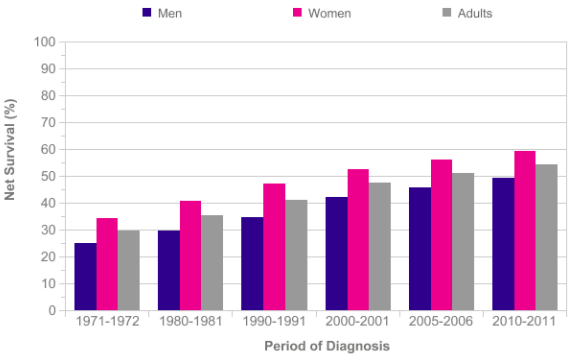

Fig. 2 Age-standardized 5-year net survival of all cancers, excluding non-melanoma skin cancer. Data collect in England and Wales[16].

Pediatric cancer

Childhood cancer is a rare event. In fact, pediatric cancer represents 1% of all cancer cases[17]. However, every year approximately 175.000 children are diagnosed with cancer worldwide[18], and pediatric cancer is still the leading cause of death by disease in children over 1 year of age in developed countries[6,19].

1. Pediatric and adult cancer

Pediatric and adult cancer differ in etiology, tumor type frequency, mutational burden, and prognosis.

1.1 Etiology

Pediatric tumors are thought to arise from embryonic developing tissues (Fig.3)[19], while adult cancers originate mostly from epithelial cells[17]. The involvement of environmental factors in pediatric cancer development is still up to discussion. In adults, it is very well documented the relation between carcinogenesis and cigarette smoke, alcohol consumption, and sun exposure. On the other hand, in children, only a few chemicals (such as dioxin, pesticides, solvents and metals) have been associated with cancer but the effect of other important carcinogens, like ionizing radiation, is still debatable[17].

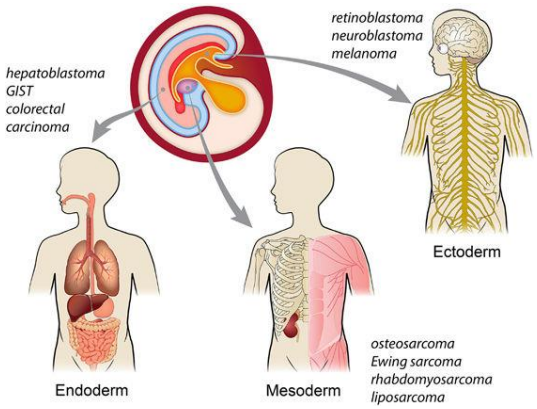

Fig. 3 Pediatric solid tumors derived from the three germ layers. Adapted from[22].

1.2 Tumor type frequency

In children, the most common tumors are acute lymphoblastic leukemia (ALL), brain tumors (medulloblastoma and neuroblastoma), lymphomas, and soft-tissue sarcomas (rhabdomyosarcoma, Ewing sarcoma, osteosarcoma, and Wilms tumor), while adults mainly suffer from malignancies of the gastrointestinal tract, lung, prostate, and breast (Fig.4)[19]. Amongst childhood cancer, there are also differences in tumor type incidence between young children and adolescents. Leukemia, followed by brain and other nervous system malignancies, is the most common cancer from the age of 1 to 14, while from the most common cancer is lymphoma, particularly Hodgkin lymphoma[6].

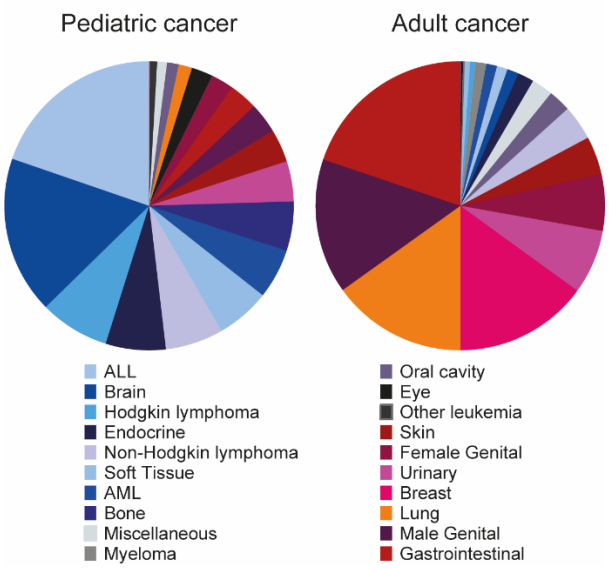

Fig. 4 Frequency of cancer subtypes diagnosed in children and adults[19].

1.3 Mutational burden

Adult solid tumors display an average of 33 to 66 somatic mutations, but melanoma or lung cancer can harbor up to 200 non-synonymous mutations. In contrast, pediatric cancers have an average of only 9.6 mutations per tumor (Fig.5). One reason for this disparity is the fact that the number of mutations is directly related to the age of the patient. Moreover, potent mutagens, such as UV light and cigarette smoke, are involved in adult but not in childhood cancer[13].

Within the same tumor type, the mutations found in adults and children can differ. Pediatric glioblastomas are rich in a histone 3 lysine 27 (H3K27) mutation, but in adults this mutation is never observed. Within child demography there are also differences. In neuroblastoma, adolescent patients present mutations in ATRX, while infants do not[19].

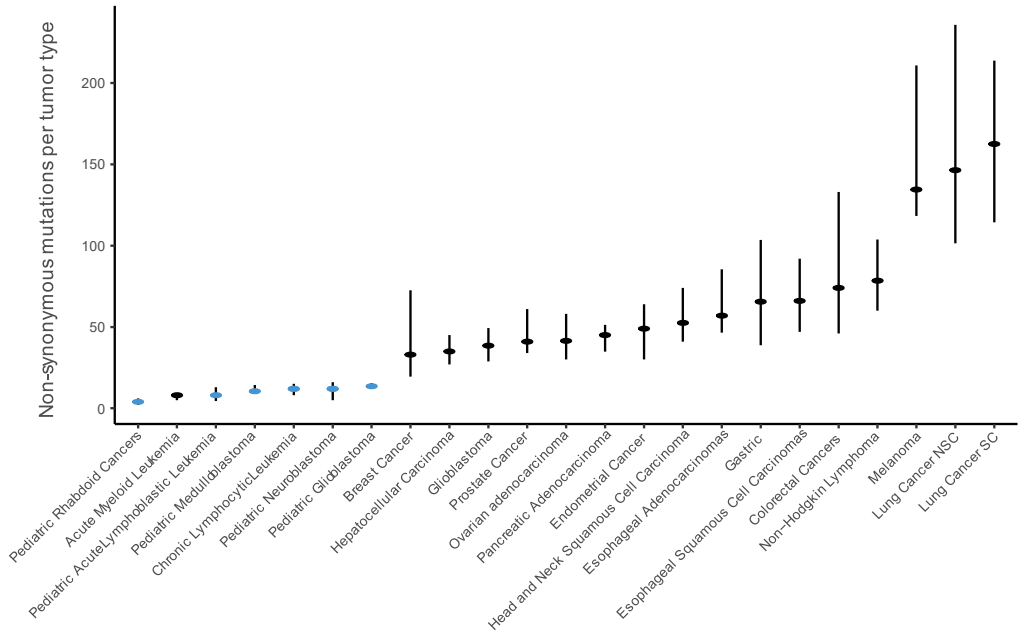

Fig. 5 Number of somatic mutations in human cancer. The plot represents the median number of non-synonymous mutations per tumor type (pediatric tumors are displayed blue). Horizontal bars indicate the 25 and 75% quartiles. Adapted from[13]. NSC – non-small cell, SC – small cell.

1.4 Prognosis

Cure rates are significantly higher in children than in adults. The 5-year relative survival rate in children diagnosed with cancer between 2007-2013 was of 83% versus only 69.2% in adults[6,20].

Since 1970, the death rates of childhood and adolescent cancers decreased 66%[6]. This change is mainly attributed to the enrollment of large numbers of patients in clinical trials, improvement of risk assessments, and targeted therapy. However, this decrease in death rate is very heterogeneous amongst tumor types, and is mainly due to a 75% reduction in non-Hodgkin lymphoma and a 60% reduction in ALL and acute myeloid leukemia (AML)[21]. In contrast, the 5-year survival rates of solid tumors have plateaued for the last 10-20 years[18]. Importantly, children who survive cancer have a higher mortality risk due to late effects of the therapy such as increased likelihood of developing subsequent malignant neoplasms (50% more risk) and cardiac or pulmonary chronic diseases. In fact, of the 80% of children that survive the first malignancy 20% will die from recurrence[17].

2. The origin of pediatric cancer

Adult cancers occur gradually by accumulation of genetic lesions. The low mutational burden observed in childhood malignancies (Fig.5) suggests a different tumorigenesis process. One possibility is that childhood cancer is caused by acute genetic changes like chromothripsis and chromosomal translocations[22]. In fact, adult cancers with fusion driving mutations bear a lower mutational burden, comparable to the one found in pediatric cancers[17].

As previously stated, pediatric malignancies arise from developing tissues when multipotent progenitor cells undergo differentiation[17,22]. The age of onset of the malignancy is possibly related to the differentiation timing of the cell of origin. Retinoblastoma is rarely diagnosed in children over 5 years of age, and the proliferation of retinal progenitor cells occurs in the first 2 trimesters of gestation. On the other hand, osteosarcoma is most commonly diagnosed in adolescents and young adults which coincides with the rapid growth of the skeletal system during puberty[22].

The establishment of mouse models that recapitulate the features of pediatric tumors is still challenging. Even when the driver mutation is known, it needs to occur in a certain timing and on the right cellular background in order to lead to cell transformation[22].

3. Chromosomal translocations, typical pediatric tumor drivers

Chromosomal translocations juxtapose two previously separated genes. The first reciprocal chromosomal translocation, the Philadelphia chromosome, was discovered in 1973. This translocation occurs between chromosomes 9 and 22 and results in the expression of the BCR-ABL fusion protein which drives chronic myeloid leukemia (CML).

Frequently, the result of a chromosomal rearrangement is a functional chimeric protein that acts as a transcription factor, driving aberrant gene expression (e.g. PAX3-FOXO1), or as a kinase, impairing signaling pathways (e.g. NPM-ALK)[17]. Fusion genes can also result in the inactivation of tumor suppressor genes (by truncation or separation of the gene from its promoter)[13] or lead to overexpression of oncogenes (by juxtaposing the oncogene to highly active regulatory sequences, e.g. IGH-MYC or TMPRSS2-ERG)[17].

Many pediatric cancers, mainly hematologic and soft-tissue malignancies, bear chromosomal translocations that drive tumorigenesis (Table 1)[9]. These fusion genes are uniquely expressed in defined tumor types, which makes them good biomarkers. The fusion EWS-FLI1 is characteristic of Ewing sarcoma, PAX3/7-FOXO1 of alveolar rhabdomyosarcoma and SSX18-SSX1/2 of synovial sarcoma[17]. Their tumor specificity and their role in tumorigenesis makes them attractive therapeutic targets. This became clear

with development of Gleevec, a BCR-ABL tyrosine kinase inhibitor, that accomplished an impressive 80% remission rate in CML[17].

Table 1 Common recurrent translocations in soft tissue tumors. Adapted from[9].

Tumor	Genetic abnormality	Fusion transcript
Ewing sarcoma	t(11;22)(q24;q12) t(21;22)(q22;q12) t(7;22)(p22;q12) t(17;22)(q12;q12) t(2;22)(q33;q12)	FLI1-EWS ERG-EWS ETV1-EWS E1AF-EWS FEV-EWS
Desmoplastic small round cell tumor	t(11;22)(p13;q12) t(11;22)(q24;q12)	WT1-EWS FLI1-EWS
Synovial sarcoma	t(X;18)(p11.23;q11) t(X;18)(p11.21;q11)	SSX1-SYT SSX2-SYT
Alveolar rhabdomyosarcoma	t(2;13)(q35;q14) t(1;13)(p36;q14)	PAX3-FKHR PAX7-FKHR
Malignant melanoma of soft part (clear cell sarcoma)	t(12;22)(q13;q12)	ATF1-EWS
Myxoid liposarcoma	t(12;16)(q13;p11) t(12;22)(q13;q12)	CHOP-TLS(FUS) CHOP-EWS
Extra skeletal myxoid chondrosarcoma	t(9;22)(q22;q12)	CHN-EWS
Dermatofibrosarcoma protuberans and Giant cell fibroblastoma	t(17;22)(q22;q13)	COL1A1-PDGFB
Congenital fibrosarcoma and Mesoblastic nephroma	t(12;15)(p13;q25)	ETV6-NTRK3

Pediatric sarcomas

Sarcomas are rare and heterogeneous malignancies that arise from bone or soft tissues such as muscle or fat[23]. They account for 1% of all cancers in adults and about 13% in children[23]. The 5-year survival rate of pediatric sarcomas is roughly of 60% but falls to 20-30% in recurrent and metastatic cases[23]. Due to their rarity, this is a group of understudied diseases.

1. Rhabdomyosarcoma

Rhabdomyosarcoma (RMS) is a group of small, round, blue-cell tumors with features of skeletal myogenesis[24,25]. While rare in adults, in children it is the most common soft-tissue sarcoma, accounting for half of all soft-tissue sarcomas and 5% of all pediatric sarcomas[24].

RMSs are mostly sporadic tumors, although a few have been associated with familial syndromes such as Li-Fraumeni syndrome (characterized by germline mutations of the tumor suppressor gene p53), Neurofibromatosis I and Costello syndrome[26].

1.1 Diagnosis

RMS is usually an asymptomatic disease. Rarely, pain on the primary tumor site is felt due to the compression of adjacent tissues[26]. Diagnosis is based on imaging techniques, such as computerized tomography (CT), magnetic resonance imaging (MRI) or positron-emission tomography (PET) scans, and confirmed by biopsy[26]. The tissue collected in the biopsies is used for histological subtyping and immunohistochemistry. Desmin, vimentin, MyoD1, Myogenin, S100, EMA, LCA, FLI1, and Mic2 are some of the biomarkers tested[27]. Ideally, FISH (fluorescence in situ hybridization) is performed to identify the presence of the fusion protein.

1.2 Histologic classification

The World Health Organization (WHO) has classified RMS in four histological subtypes: embryonal, alveolar, pleomorphic, and sclerosing/spindle cell (Fig.6)[28].

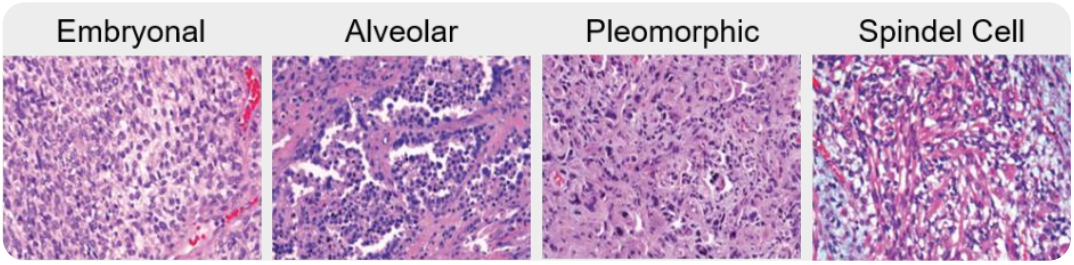

Fig. 6 RMS subtypes based on histology. Adapted from[29].

Embryonal RMS (eRMS) resembles immature skeletal muscle and is the most prevalent subtype, accounting for 60% of all RMS cases. It mainly occurs in the head and neck, genitourinary tract, and retroperitoneum[25]. ERMS has a bimodal age distribution with the first peak of incidence occurring in children under 5 years and the second one in adolescents[24].

The alveolar subtype (aRMS), as the name suggests, histologically resembles lung alveoli. It accounts for 20% of all RMS cases and mainly takes place in the trunk and extremities[24]. Although it can affect all ages, it is more common in adolescents and young adults[24] unlike the pleomorphic and the sclerosing/spindle variants which mainly affect adults[26,28,29]. Pleomorphic RMS is an aggressive neoplasm that typically arises in deep soft tissues of the extremities of patients older than 45 years[26]. These neoplasms have complex karyotypes and are highly aggressive[29].

Finally, sclerosing/spindle cell RMS is generally found in the extremities and head and neck region in adults, while in children is mainly paratesticular. Similarly to the other subtypes, sclerosing/spindle cell tumors are more aggressive in adults than in children[28,29].

1.3 Molecular classification

The two main RMS subtypes, eRMS and aRMS, have significantly different genomes. ERMS has a complex karyotype[24], while aRMS presents very few mutations[30] and, in 70% of the cases, is characterized by a specific translocation, resulting from a rearrangement of the PAX gene (Fig.7A)[31]. The presence of this fusion gene deeply affects tumor biology. In fact, aRMS tumors negative for the PAX fusion (circa 20% of all aRMS cases) are more similar to eRMS on the transcriptome level[30,32–34] and in mutation burden (Fig.7A). In fact, fusion positive aRMS carry an average of 6.4 non-synonymous somatic mutations, while eRMS and fusion negative aRMS have approximately 17.8[30]. The fusion status also defines the aggressiveness of the tumor. At the time of diagnosis, 46% of fusion positive aRMS, against only 17% for fusion negative, bear detectable metastasis[35]. Indeed, expression of the PAX fusion in the eRMS cell line RD increases cell growth and produces more invasive mouse xenografts[36]. Moreover, the prognosis of fusion negative aRMS is significantly better than the one of fusion positive and indistinguishable from eRMS (Fig.7B)[31].

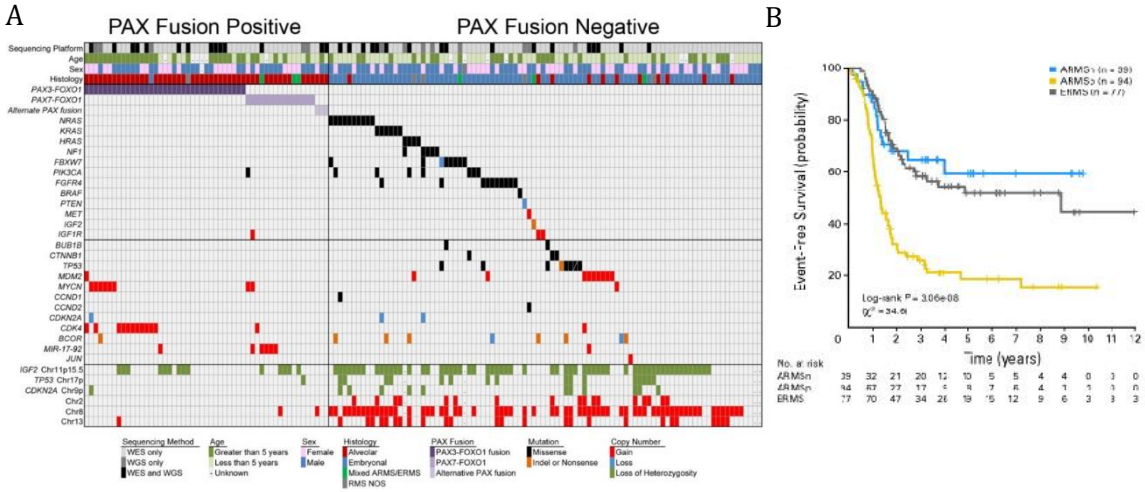

Fig. 7 Clinical and molecular features of FP- and FN-RMS. (A) The genomic landscape of pediatric RMS. Adapted from[37]. (B) Kaplan-Meier curves showing the event-free survival for patients with fusion positive aRMS (ARMSp), fusion negative aRMS (ARMSn) and eRMS (ERMS). Adapted from[31].

Together, these findings suggest that the fusion status represents the tumor types better than histology in terms of genetics, molecular, and clinical features. Therefore, I will refer to fusion positive (FP-RMS) or fusion negative (FN-RMS) rhabdomyosarcoma from now on.

1.4 Mutational burden

Recent large-scale sequencing studies of primary RMS tumors revealed a low overall mutation rate (0.31 protein-coding mutations/Mb)[37].

FN-RMS commonly presents mutations in the Ras pathway (Fig.7A), the receptor tyrosine kinase FGFR4, the catalytic component of the phosphoinositide-3 kinase complex (PIK3CA)[30], neurofibromin (NF1), β-catenin (CTNNB1), and BCOR[29]. Frequent mutations in cell-cycle genes (FBXW7 and E3 ubiquitin ligase) and TP53, or focal losses of TP53[30], as well as focal amplification of the p53 modulator MDM2, are also found[37]. Besides local mutations, loss-of-heterozygosity (LOH) is recurrently observed in FN-RMS. LOH mainly affects the 11p15.5 region, which includes the IGF2, H19, and CDKN1C genes[24,37].

In FP-RMS, mutations occur at the extremely low rate of 0.1 protein-coding mutations/Mb[37]. The region 12q13-15, which contains the cell cycle-dependent kinase CDK4 and MDM2, is amplified in 28-56% of the cases and commonly associated with a worse prognosis[25,30]. Another important amplified region, in 14-19% of the cases, is 2p24 which contains the proto-oncogene MYCN[25]. MYCN is a phosphoprotein that functions as a transcription factor and is involved in the regulation of cell proliferation and differentiation[38]. In FP-RMS, MYCN overexpression or copy number gain is associated with poor outcome[38]. Finally, an amplification of 15q25-26 is also observed and leads to the overexpression of the insulin-like growth factor type receptor I (IGF1R)[39]. Additionally, loss-of-imprinting (LOI) causes IGF2 overexpression in 50% of FP-RMS[25].

1.5 Therapy

The standard of care for RMS is a multimodal treatment based on chemotherapy and surgery, with or without radiotherapy. Chemotherapy comprises a three-drug regimen of vincristine, actinomycin D and cyclophosphamide (VAC). Radiotherapy is used in patients with unresectable disease or residual lymph node disease after surgery, and for patients of the alveolar subtype[26].

This treatment strategy was implemented in the 1970s and led to a 5-year survival close of 61% in the 1990s, which has not improved since then[35]. In addition, this therapeutic scheme is not effective for high-risk patients or patients with recurrent disease, who maintain a 5-year overall survival of 30% and 17% respectively[37].

Severe late side effects have been reported for this treatment modality. Alkylating agents, like cyclophosphamide, are associated with secondary malignancies and sterility, and radiotherapy can cause facial asymmetry and growth deficiency[26]. Hence, new targeted therapies with higher efficacy in high-risk patients and less side effects are urgently needed.

1.6 Prognosis

Classically, RMS prognosis was defined based on the site of onset, histological variant, presence of metastasis, tumor size, and age of diagnosis[26]. In addition to these factors, today we are also aware of the role of fusion status in RMS prognosis (Fig.7B)[40].

Head and neck, genitourinary tract, and bile ducts were classified as sites of favorable outcome. The alveolar subtype has a worse prognosis than the embryonal since it presents frequently metastasis at the time of diagnosis and has a poorer response to therapy. In fact, eRMS has a 5-year overall survival of 73%, while the one of aRMS only reaches 48%, dropping to 10-30% in case of metastasis. The main metastasis sites are the lungs, bone marrow, lymph nodes, and bones[24].

Adult RMS has significantly worse prognosis, with a 5–year overall of only 27% versus 61% in children. Also adolescents have a less favorable prognosis than children diagnosed between the age of 1 to 10[26].

2. Fusion positive rhabdomyosarcoma

2.1 The fusion gene

To date, 6 different PAX fusion genes have been identified in FP-RMS (Fig.8)[25,32,37]. PAX3-FOXO1, the most common rearrangement, is present in roughly 60% of all aRMS cases followed by PAX7-FOXO1, which accounts for other 20%[25]. The remaining fusions are present in roughly 10% of all aRMS cases[25].

The fusions merge the N-terminal DNA-binding paired and homeodomains of PAX3/7 in-frame with the transactivation domains of either the forkhead box O4 (AFX), the nuclear receptor coactivator (NCOA1/2), the chromatin remodeling gene (INO80D), or, more commonly, the forkhead box O1 protein (FKHR or FOXO1)[37].

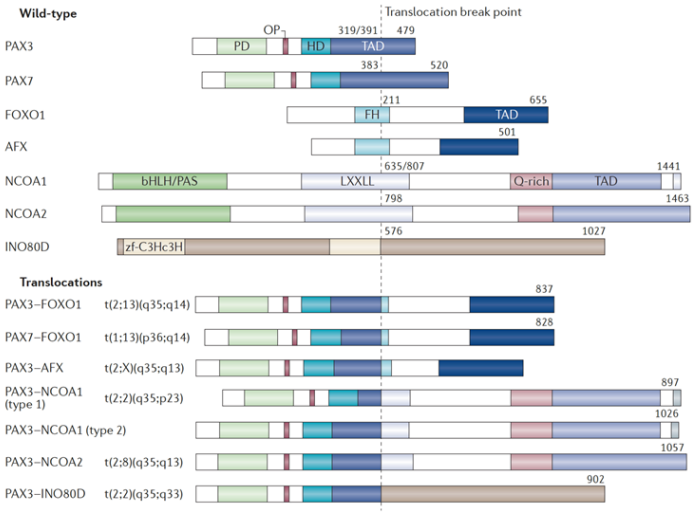

Fig. 8 Fusion proteins associated with FP-RMS and their respective wildtype counterparts. Adapted from[29].

2.2 PAX3/7-FOXO1

Despite the similarity between the fusion genes, PAX3-FOXO1 positive patients have a poorer prognosis than the ones carrying the PAX7-FOXO1 fusion[40]. In metastatic disease, the event-free survival at 4 years is 8% and 75% for PAX3- and PAX7-FOXO1, respectively[39]. In addition, PAX7-FOXO1 tumors localize more often to favorable sites and are typically present in younger patients[41]. Nevertheless, expression data have shown no difference in the repertoire of downstream targets of PAX3- or PAX7-FOXO1[34].

The fusion gene is the main driver of FP-RMS tumorigenesis. Once expressed, the fusion protein acts as a transcription factor and drives aberrant gene expression leading to cell transformation. In fact, overexpression of the fusion is able to transform mouse (NIH3T3)[42] and chicken embryo fibroblasts[43]. The tumor cells then become addicted to continuous expression of the fusion protein[24]. Antisense oligonucleotides targeting PAX3-FOXO1 drastically decrease FP-RMS cell viability[38].

PAX3-FOXO1 transformation activity is dependent not only on the DNA binding activity of the Pax gene but also on the protein-protein interactions mediated by FOXO1[42]. Ectopic expression of a fusion between the DNA-binding domains of PAX3 with the transcription repressor, KRABB, leads to downregulation of PAX3-FOXO1 targets and consequent reduction of cell growth in vitro and tumor formation in vivo[44].

In FP-RMS, PAX3/7-FOXO1 proteins are highly expressed and unresponsive to transcriptional repressors. The overexpression of PAX3-FOXO1 results from an increase in its transcription rate, while the one of PAX7-FOXO1 originates from the amplification of the gene[45]. Both fusions, unlike wildtype FOXO1, are resistant to AKT-dependent regulation of subcellular localization and are exclusively located in the nucleus[38]. Additionally, hDAXX is able to bind both wildtype PAX3 and the fusion, but PAX3-FOXO1 is resistant to its transcription repression[46].

2.3 PAX and FOXO gene families

The PAX (paired box) family consists of 9 transcription factors (TFs) which control organogenesis. This family is defined by the conserved paired domain which is composed of 128 amino acids and recognizes the TCACGC/G DNA sequence. PAX2-8 possess a second DNA binding domain, the homeodomain, that binds to the consensus sequence TAAT-ATTA. Paired and homeodomain can bind cooperatively (consensus sequence AATTA-GTCACGC) or independently. Besides PAX2 and 6, the PAX family also contains an octapeptide motif between the paired and homeodomain (Fig.8). This 8 amino acid sequence functions as a transcriptional repressor, rendering PAX proteins as transcription activators or repressors depending on the context and cofactors[41].

PAX3 and 7 are involved in the development of the nervous system and skeletal muscle[39]. Both commit cells to differentiation but at the same time inhibit terminal differentiation and sustain proliferation during migration. Once the cells reach their final destination, the PAX genes are downregulated and differentiation is completed[45]. Despite their similarity in structure, PAX3/7 have different roles in development. In somite compartments, Pax3 expression is activated before Pax7, but Pax7 persists longer. In adult mice, Pax7, but not Pax3, is expressed in most myogenic satellite cells[38].

FOXO1 belongs to the family of the forkhead transcription factors and regulates metabolism, cell proliferation and apoptosis[38]. Regulation of apoptosis is mediated by the binding of FOXO1 to the insulin responsive elements (IREs) and the regulation of proapoptotic factors[44].

FOXO1 is not detectably expressed in FP-RMS, due to the loss of one copy caused by the translocation and the altered stability of the second one[38,45]. Overexpression of FOXO1 in FP-RMS, but not in FN-RMS, leads to cell cycle arrest and apoptosis by expression of the protease caspase-3, which is consistent with the role of FOXO1 as a tumor suppressor[38].

2.4 The fusion gene signature

Initial attempts to identify PAX3-FOXO1 target genes relied on the ectopic expression of the fusion in different systems (e.g. the osteosarcoma cell line SaOS-2[47] and FN-RMS cell line RD[33]). These studies showed that the fusion gene mainly acts as a transcriptional activator and that it not only regulates PAX3 target genes, by binding to the same DNA sequence[48], but also contains its own targets. Interestingly, the constitutive expression of PAX3/7 in the skeletal muscle lineage results in developmental defects but not in tumor formation[25], which suggests that the target genes of wildtype PAX3/7 are not the ones driving tumorigenesis.

Gene ontology analysis showed that PAX3-FOXO1 activated genes control cell adhesion, proliferation, programmed cell death and neurogenesis, while the ones downregulated are involved in muscle development (Table 2)[33].

PAX3-FOXO1 activates IGFR1 expression[24], which, upon ligand binding, activates many downstream signaling pathways (MAPK, PI3K and mTOR) that contribute to cell survival[49]. Additionally, cell proliferation is promoted by the inactivation of the cyclin-dependent kinase inhibitor 2A (CDKN2A)[39] and massive upregulation of several receptor tyrosine kinases (such as FGFR4 and ALK)[37]. Another important PAX3-FOXO1 target involved in cell proliferation is the transcription factor AP2β (TFAP2β). In fact, depletion of TFAP2β causes FP-RMS cell death by apoptosis[25]. The fusion gene also activates the expression of the anti-

apoptotic protein BCL-xL, which protects FP-RMS from cell death. Indeed, expression of BCL-xL rescues the cell death phenotype upon P3F depletion[50].

The invasive properties of FP-RMS have been associated with the upregulation MET, CXCR4 and CNR1[25]. MET is a tyrosine kinase receptor for the hepatocyte growth/scatter factor (HGF/SF). During skeletal muscle development, HGF/SF is downstream of PAX3[25] and promotes myoblast proliferation[39]. In FP-RMS, MET levels are correlated with a poorer prognosis and its depletion leads to apoptosis[25]. CXCR4 is a G-protein coupled chemokine surface receptor, normally expressed in satellite cells within skeletal muscle. In FP-RMS, CXCR4 induces chemotaxis, helping cell migration and invasion[25]. The cannabinoid receptor 1 (CNR1) has been associated with the formation of lung metastasis[25].

MYCN and CDH3, both direct targets of the fusion, contribute to cell transformation. CDH3 (P-Cadherin) expression maintains C2C12 myoblasts in a proliferative state and inhibits their differentiation[25]. MYCN is a transcription factor of the neural development and when co-expressed with PAX3-FOXO1 in the mouse NIH3T3 fibroblasts greatly improves the transformation activity of the fusion[51].

PAX3-FOXO1 keeps FP-RMS in an undifferentiated state while conferring it myogenic traits. The expression of the fusion gene in murine fibroblasts and mesenchymal stem cells activates the myogenic program but its expression in the murine myoblasts C2C12 inhibits terminal differentiation[24]. PAX3-FOXO1 induces MYOD and myogenin (MYOG) expression, committing cells to the myogenic program[24]. However, simultaneously, the fusion gene prevents MYOD1 activity[24], thus inhibiting terminal differentiation, by keeping it phosphorylated and hence inactive[52].

Table 2 PAX3-FOXO1 target genes.

Function	Target genes	Expression in FP-RMS
Proliferation	IGFR1	Upregulated
	FGFR4	Upregulated
	ALK	Upregulated
	MET	Upregulated
	TFAP2β	Upregulated
	BCL-xL	Upregulated
	CDKN2A	Downregulated
Invasion	MET	Upregulated
	CXCR4	Upregulated
	CNR1	Upregulated
Transformation	MYCN	Upregulated
	CDH3	Upregulated
Myogenic differentiation	MYOD	Upregulated
	MYOG	Upregulated

2.5 Cell of origin

Due to the myogenic features present in FP-RMS, it has been postulated that it may arise from muscle precursor cells. Nevertheless, rhabdomyosarcomas are found in organs that lack skeletal muscle, such as prostate, urinary bladder, and gallbladder, which suggests a highly plastic cell of origin. To date, mesenchymal stem cells (MSC) and satellite cells have been proposed as possible cells of origin of FP-RMS. MSCs have the ability to self-renew but are also committed to a differentiation program that originates either adipocytes, osteoblasts or skeletal myocytes (Fig.9). Interestingly, Pax3 commits MSCs to the muscle lineage. Satellite cells are the myogenic progenitors of postnatal muscle. These cells differentiate into myoblasts during postnatal growth or in response to damage[45]. PAX3-FOXO1 expression in mice satellite cells does not lead to tumor development but to a smaller number of postnatal myoblasts. In MSCs, the expression of PAX3-FOXO1 induces the expression of MYOD and MYOG[45]. These results point to MSCs as the possible cell of origin of FP-RMS. However, MSCs form tumors if not only PAX3-FOXO1 is expressed but also if the p53 pathway is inactivated. This is contradictory to sequencing studies which have shown that, occasionally, the fusion protein is the only somatic genetic alteration encountered[30]. Hence, the origin of FP-RMS is still questionable.

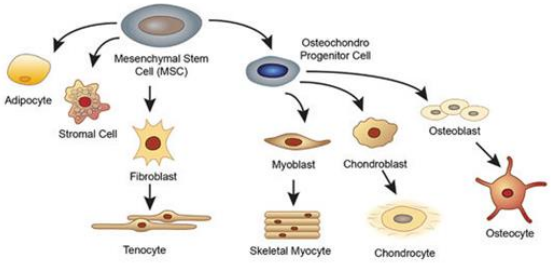

Fig. 9 Differentiation routes of mesenchymal stem cells differentiation. Adapted from[22].

2.6 Targeted Therapy

FP-RMS is dependent on continuous expression of its unique fusion protein and on its oncogenic gene expression program. Thus, interfering with PAX gene rearrangements, directly or indirectly, provides means to specifically target FP-RMS.

One strategy to hamper PAX3-FOXO1 activity is to act on its target genes. For example, blocking IGFR1 impairs FP-RMS cell growth in vitro[53]. Unfortunately, clinical trials with the human immunoglobulin G monoclonal antibody against IGF1R, Cixutumumab[37], showed no durable objective clinical response[35]. Pre-clinical studies have also shown the potential of FGFR4 inhibition, by the tyrosine kinase inhibitor ponatinib[37]. This drug suppresses STAT3, a downstream target of FGFR4[53], and inhibits RMS cell lines proliferation in vitro and tumor growth in mouse xenographs[37].

Direct targeting of PAX3-FOXO1 is more challenging due to its loose structure. Hence, strategies targeting the fusion stability have been developed. CDK4 inhibition abrogates the phosphorylation and the nuclear localization of PAX3-FOXO1. Treatment of FP-RMS cells with the CDK4/6 inhibitor palbociclib leads to G1 arrest and the expression of muscle-

specific markers[37]. PLK1 inhibitors increase ubiquitylation and degradation of PAX3-FOXO1, and promote tumor regression in xenograft mouse models[29]. Furthermore, the HDAC inhibitor entinostat decreases PAX3-FOXO1 mRNA levels, and, consequently, reduces tumor growth[54].

3. Ewing Sarcoma

Ewing Sarcoma (ES) is the second most common bone sarcoma in childhood and adolescence[55], and was first described in 1866 as a small-round cell tumor[56]. In 1921, James Ewing separated this tumor entity from the osteogenic sarcoma and renamed it as "diffuse endothelioma of the bone"[57]. In the 1980s, a chromosome analysis detected the characteristic t(11;22)(q24;q12) translocation in ES cell lines, which later in 1992[58] was found to produce the chimeric protein and tumor driver EWS-FLI1[56].

3.1 Disease prevalence

The peak of ES incidence occurs between the ages of 15-24[59] and is rarely diagnosed after the age of 30[55]. Nevertheless, to date, the oldest patient diagnosed with ES was 77 years old[60]. The average incidence of ES in adolescents and young adults is 2.6 per million, with a higher predominance in the male population (1,5:1), and rarely described in the black population[55,56,61]. The low susceptibility of the African population can be explained by their distribution of GGAA microsatellites. EWS-FLI1 has higher activity when GGAA microsatellites have a length of 20-30 repeats which is more commonly found in the Caucasian population compared to the African[62,63].

ES occurs mainly in bones (Fig.10) and is generally found in the diaphyseal portion of flat bones[64]. The most commonly affected bones are pelvis, ribs, spine, and femur[55,65]. However, in 30% of the cases, this tumor is found in soft tissue[59]. Extra skeletal regions include kidney, breast, gastrointestinal tract, prostate, endometrium, lung, adrenal gland, and meninges[56,64]. The prognosis for these locations is unfavorable in the first 2 years but afterwards the outcomes are significantly better compared to skeletal ES[66]. Primary metastasis in lungs, pleural space, bone, and bone marrow are detectable in 25% of patients at the time of diagnosis[64].

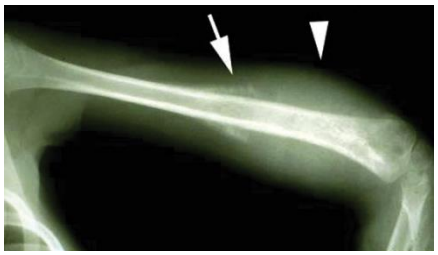

Fig. 10 Radiograph shows a ES mass in the metadiaphyseal region of the distal humerus. Adapted from[65].

3.2 Diagnosis

Pain at rest is the most common symptom present in ES patients[64,67]. It is usually locoregional, intermittent, and often erroneously attributed to trauma[67,68]. Other symptoms include swelling, venous dilation, palpable mass formation, paresthesia, and fever in the advanced stages[56,64].

The initial diagnosis is made through imaging techniques, such as CT and MRI, but definitive diagnosis is achieved by biopsy[56,64]. Provided enough biopsy material is available, conventional histology, immunohistochemistry, and molecular biology analysis are performed. Cytoplasmic glycogen and CD99 are usual ES makers, although not specific[65,68]. The identification of the characteristic EWS-FLI1 and EWS-ERG fusion proteins confirms the ES diagnosis[69]. In the clinics, the detection of the fusion protein is accomplished by reverse transcription polymerase chain reaction (RT-PCR) or FISH (Fig.11A)[56]. Nevertheless, a negative RT-PCR does not exclude ES since EWS-FLI1 and EWS-ERG are not the only fusion genes present in ES[65].

3.3 Histology

ES represents a group of tumors with varying degrees of differentiation and can be divided into classical ES, primitive neuroectodermal tumor (PNET), and atypical ES[65]. Classical ES (Fig.11B), as described by James Ewing, is composed of "broad sheets of small polyhedral cells with pale cytoplasm, small hyperchromatic nuclei, well-defined cell borders, and complete absence of intercellular material"[57]. PNET are ES tumors with evident signs of neural differentiation such as Homer Wright rosettes or the presence of specific markers (neuron-specific enolase, S-100 protein and Leu-7)[64,65,68]. Atypical ES (large cell ES) refers to tumors with features that deviate from those of classical ES, such as nuclear enlargement and pleomorphism, irregularity of nuclear membrane, and prominent nucleoli.

The new WHO report considers the distinction between ES and PNET obsolete[70] since both express the oncogenic fusion protein and present no clinical differences.

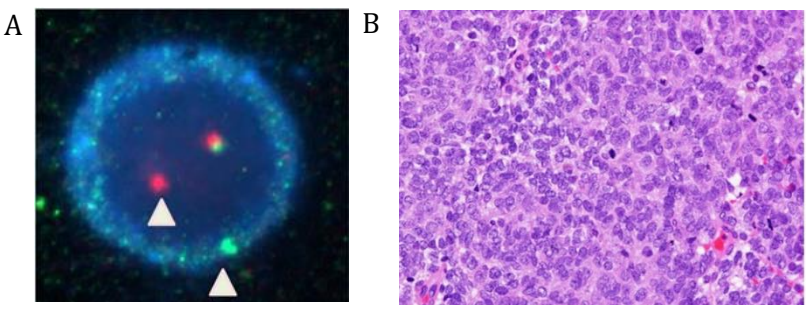

Fig. 11 (A) FISH detection of the EWS-FLI1 fusion protein. (B) Classical ES sarcoma histology. Adapted from[56].

3.4 Mutational burden

Besides the distinctive translocation, ES is characterized by a relatively simple karyotype (0.15 mutations/Mb[71]) with only a few numerical and structural aberrations (Fig.12)[72].

Structural changes of chromosomes 1 and 16 are present in roughly 20% of the tumors, which result in a gain of 1q and a loss of 16q, and the formation of the derivative chromosome der(1;16)[64]. Gain of chromosome 1q is associated with the overexpression of CTD2 (cell division cycle protein 2), a protein involved in DNA damage repair[73]. Trisomy 8 is present in nearly half[72] of the cases and trisomy 12 in a third[64]. Chromosome 8 gain leads to an overexpression of the oncogene MYC, while chromosome 12 gain is associated with high expression of NANOG, CDK4, ERBB3, GLI1, and MDM2[73]. Homologous deletion of the chromosomal region 9p21, containing ink4A (which encodes for the well-known cell cycle regulator CDKN2A), is present in 25-28% of ES cases[64,72]. LOH at 17p13, followed by mutation of the remaining p53 allele, occurs in less than 10% of the cases[64].

Mutations in TP53, STAG2 and CDKN2A are recurrent[71,74,75]. Mutations in TP53 occur in 6% of the cases and are usually accompanied by STAG2 mutations[71,72,75]. The presence of both TP53 and STAG2 mutations is correlated with unfavorable prognosis[75]. STAG2 mutations, which are associated with loss of STAG2 expression, are present in more than 15% of ES cases[72]. Interestingly, 88% of patients with STAG2 mutations present metastases at the time of diagnosis[72], suggesting an association between STAG2 status and disease dissemination. CDKN2A is mutated in 12% of all ES cases and, interestingly, STAG2 mutations and CDKN2A deletions are mutually exclusive[75].

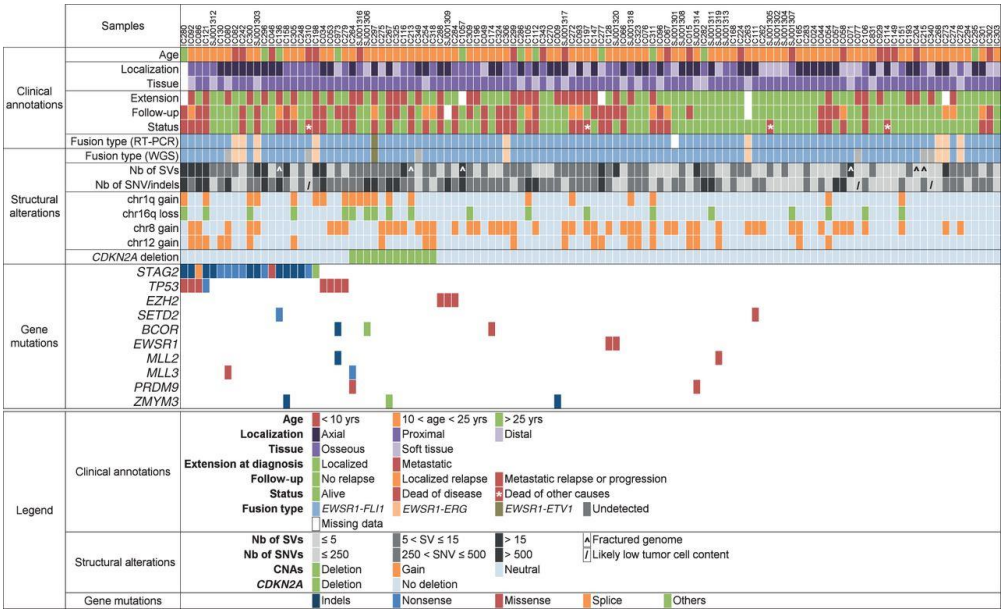

Fig. 12 Mutation landscape of ES. Adapted from[75].

3.5 The fusion genes

Ewing sarcoma is characterized by recurrent translocations that occur between members of the TET and ETS families of transcription factors. These chromosomal translocations originate different fusion genes that act as aberrant TFs and are associated with tumor development[76].

EWS-FLI is the most common fusion gene, being present in 85% of all ES cases[76]. This oncoprotein fuses the DNA-binding domain of FLI1 (friend leukemia virus integration site 1; the ETS family member) in-frame with the N-terminal domain of EWSR1 (the TET family member), under the control of the EWSR1 promoter[58]. In ES, EWS is also found fused to other ETS family members (Fig.13) such as ERG (ETS-related gene)[77], ETV1 (ETS-variant gene 1)[78], ETV4 (ETS-variant gene 4, also called E1AF - adenovirus E1A enhancer-binding protein)[79] and FEV (fifth Ewing sarcoma variant)[80]. EWS-ERG, the result of the t(21;22)(q22;q12) translocation, is the second most common rearrangement found in ES, accounting for 10% of the cases. ERG shares 68% of overall amino acid sequence with FLI1 and studies have shown no clinical difference between ES tumors bearing EWS-FLI1 or EWS-ERG fusion genes[76].

All EWS-ETS fusion proteins have similar structure and function. The amino-terminal transactivation domain of EWSR1 is fused to the ETS DNA-binding domain[76,81]. These fusions are all able to form tumors in mice when overexpressed in murine fibroblasts NIH3T3[82]. Moreover, overexpression of a dominant negative mutant of EWS-ETV1 can inhibit NIH3T3 cell transformation by EWS-FLI1[83], which proves the redundancy of EWS-ETS fusions. The transformation activity of the fusion protein is dependent on EWSR1 and both the DNA-binding domain of FLI1[84] and the COOH terminus of FLI1[85].

Infrequently, non-EWS fusions have been identified in ES. In these cases, the fusion occurs between another TET family member named TLS (also called FUS) and the ETS members ERG[86] or FEV[87]. These oncoproteins are thought to also recapitulate the oncogenic role of EWS-FLI[76]. Non-TET-ETS rearrangements have also been identified in small round cell sarcomas which histologically resemble ES and are categorized by WHO as Ewing-like sarcomas[65].

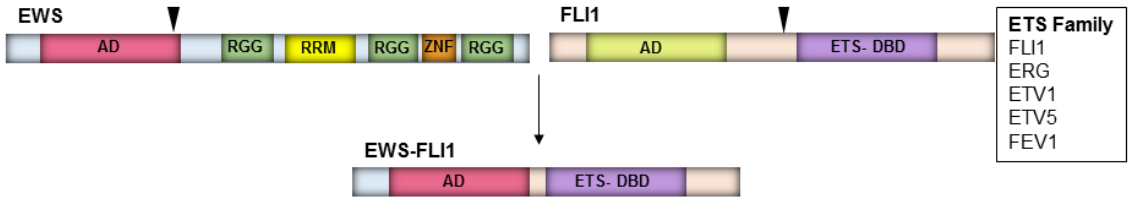

Fig. 13 Schematic representation of the EWS-FLI1 translocation. AD - Activation domain; RGG - arginine-glycine-glycine-rich region; RRM - RNA recognition motif; ZNF - zinc finger domain; ETS-DBD - ETS DNA binding domain. Arrows indicate the regions where the translocation occurs. Adapted from[81].

3.6 EWS-FLI1

From the reciprocal translocation t(11;22)(q24;q12), only the EWS-FLI1 product is expressed[58,64]. The amino-terminal domain of EWS contains the EAD (EWS activation domain)[76] which is composed of several repeats of serine-tyrosine-glycine-glutamine and is a stronger transcriptional activator than transactivation domain of FLI1[88]. The carboxy-terminal of FLI1 contains the ETS DNA-binding domain and recognizes purine-rich sequences with the GGAA/T core motif[76].

EWS-FLI, rather than a single entity, is a group of highly related isoforms with different intronic breakpoints in the EWSR1 and FLI1 genes (Fig.14). The most common fusion is called type I and consists of EWSR1 exons 1-7 and FLI1 exons 6-10[76]. The relevance of the different fusions in the pathology of the disease is still unclear.

Similarly to PAX fusions in FP-RMS, EWS-FLI1 is essential for ES. Depletion of EWS-FLI1 in ES cells inhibits cell proliferation in vitro, by apoptosis or cell cycle retention in G1[89], and tumor growth in vivo[90,91].

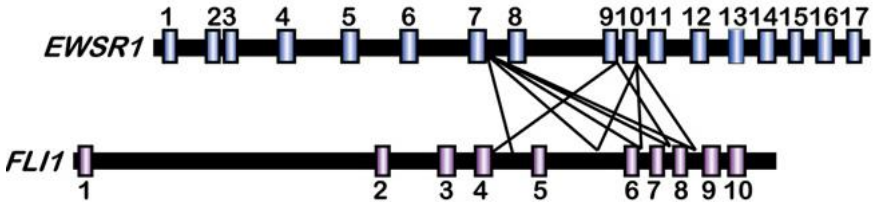

Fig. 14 Genomic structure of the EWSR1 and FLI1 genes. Exons are numbered and represented by rectangles and the connecting lines represent the breakpoints in the EWSR1 and FLI1 genes found in EWS-FLI1 translocations. Adapted from[76].

3.7 EWSR1 and FLI1 gene families

The EWSR1 gene encodes the EWS protein, a member of the TET family which also includes TLS and TAF15[76]. EWS plays a role in meiotic cell division, mitotic spindle formation, DNA repair, and ageing[92]. In its C-terminal, the EWS protein contains an RNA-binding domain which is important for RNA metabolism. EWS associates with members of the transcriptional machinery, including RNA polymerase II, TFIID and CBP/300, which suggests an active role in transcription regulation[76].

FLI1 encodes the FLI1 protein, a member of the ETS (E-26 transformation specific) family of transcription factors[76]. This family is characterized by a highly conserved winged helix-loop-helix DNA-binding domain, known as the ETS domain[76]. All 27 members of the ETS family bind to the core GGAA/T motif[93]. ETS members function as signal-dependent transcriptional regulators of cellular differentiation and proliferation[93]. Besides ES, other translocations involving ETS family members have been involved in other cancers such as the TMPRSS2-ERG fusion in prostate cancer[93].

3.8 Fusion gene signature

Two main approaches were used to identify EWS-FLI1 target genes. Either the fusion protein was ectopically expressed in a non-ES cell type (fibroblasts[94], FN-RMS cells[95], and MSCs[96]) or EWS-FLI1 was silenced in ES cell lines by RNAi[97–100]. These studies revealed that EWS-FLI1 can both activate and repress gene expression. A meta-analysis performed on the previously mentioned studies identified a core signature composed of 503 activated genes, and 293 repressed[101]. The activated core included NR0B1[98], NKX2.2[97], PTPN1[102], ID2[103], CAV1[99], and MYC[94]. Among the downregulated signature there was the insulin-like growth factor binding protein 3 (IGFBP3)[100,101].

Regarding molecular functions, genes upregulated by the fusion are implicated in oncogenic transformation (NR0B1[98], NKX2.2[97], GLI1[104], c-Myc[95], CAV1[99]), cell proliferation (PTPN1[102], ID2, CCND1 and CCNE[105]), drug-resistance (GSTM4[106]), immortalization (hTERT[107]), angiogenesis (VEGF[94]) and maintenance of pluripotency (EZH2[96]). On the other hand, genes downregulated by EWS-FLI1 are involved in cell cycle progression (FOXO1[108], p27 and p21[105,109]), evasion of apoptosis (IGFBP3[89]), evasion of growth inhibition (TGFBRII[110]), apoptosis regulation (PHLDA1[111]) and cell transformation (LOX[112]).

EWS-FLI-1 also induces the expression of neural cell markers including NYP1R, GRP, MSX1, EGR2, NKX2-2, NGFR, CITED2, CDH11, and MAPT[95].

3.9 Treatment

Current treatment strategies include chemotherapy, surgery and radiotherapy[56]. Chemotherapeutic protocols consist of a combination of vincristine, actinomycin-D, cyclophosphamide, and doxorubicin (VACD) with or without ifosfamide and etoposide (VACD-IE)[56]. Chemotherapy is applied preoperative to eradicate micro metastases and reduce the tumor volume. Local lesions are surgically removed, and radiotherapy follows with adjuvant chemotherapy[56,64]. If the tumor is inoperable, radiotherapy alone is employed[68]. Radiotherapy is applied at 40-45 Gy doses for microscopic and 50-60Gy for macroscopic disease[68]. Relapses occur in the first 3 years after therapy[68] and there is still no standard treatment available for those[56].

Late effects of therapy comprise poor local orthopedic outcome, caused by surgery and radiotherapy, and systemic effects from radio and/or chemotherapy. Radio and chemotherapy can cause secondary malignancies, being leukemia the most common. The usage of anthracyclines, like doxorubicin, induces a dose-related cardiomyopathy. Alkylating agents (cyclophosphamide and ifosfamide) are associated with infertility and ifosfamide can also cause persistent renal tubular electrolyte loss and decrease in glomerular function[64].

3.10 Targeted therapy

Like in FP-RMS, the unique presence of the fusion protein and the tumor susceptibility to its continuous expression makes EWS-FLI1 an ideal target for therapy. Therefore, new therapies have been developed to influence the activity of the fusion by either targeting its binding partners, its target genes or by influencing it directly.

PARP1 was shown to interact with both EWS-FLI1 and EWS-ERG, which linked these to DNA damage reponse[113]. Moreover, ES is defective in many DNA-damage repair genes, such as BRCA1, GEN1, and ATM[114]. Naturally then, ES cells are sensitive to PARP inhibitors, like olaparib, and this sensitivity is increased by the combination with DNA-damaging drugs, like irinotecan or temozolomide[113,114]. Unfortunately, in a phase II clinical trial, olaparib failed to produce a significant response[115]. EWS-FLI1 also interacts with the RNA helicase A, an integral component of protein complexes that regulate transcription and splicing[116]. This interaction can be blocked by the small molecule YK-4-279[117], which induces apoptosis in ES cells and reduces the growth of mouse tumor orthotopic xenografts[117].

The IGF pathway is active in ES by upregulation of IGF1[118] and downregulation of IGFBP3, which blocks the interaction between IGF1 and its receptor[100]. Inhibition of this pathway through small molecule or monoclonal antibodies induces cell death and reduces ES tumor growth[119,120]. Clinical trials have shown that inhibitors of the IGF pathway can have a valuable effect in a selection of ES patients[121].

Interestingly, the chemotherapeutic actinomycin D can directly influence EWS-FLI1. This DNA-binding molecule specifically removes EWS-FLI1 from the chromatin and affects a wide range of the fusion target genes[122]. Sadly, due to its toxicity, it is not possible to achieve in humans the concentration necessary of actinomycin D for this effect[122]. Nevertheless, a similar effect on EWS-FLI1 target genes was observed with another chemotherapeutical drug, trabectedin[123].

3.11 Prognosis

The current treatment strategies have improved the 5-year overall survival of ES from 10% to 60%, in localized disease[124], and to 20-30% for metastatic disease[68]. Bone metastases confer poorer prognosis than the lung ones[64]. Age of diagnosis also plays a role in prognosis. Patients younger than 15 years have a 5-year overall survival of over 65% while it drops to 50% for older patients[55]. Primary tumor location in the appendicular skeleton and tumor size ≤8 cm confer significant survival benefit, while race and neuroectodermal differentiation do not add prognostic value[61,65]. Despite the variety of breakpoints of EWS-ETS fusions, there is no significant prognostic value of the fusion type on the risk of progression or relapse[124].

3.12 Cell of origin

James Ewing first predicted an endothelial origin for ES due to its similarity to angio-endothelioma[57]. Since then, many cell types have been considered as possible cradles of ES, being neural crest cells and mesenchymal stem cells the most promising candidates[125].

The neural crest cell hypothesis is supported by the presence of neuroectodermal markers and neuronal differentiation characteristics in some ES tumors[65,126]. Nevertheless, the occurrence of neuronal related morphology in ES varies from absent to prominent[127]. On the other hand, ES cells exposed to differentiation-inducing agents clearly undergo neuronal differentiation[128] and neural crest stem cells are permissive to the expression of ectopic EWS-FLI1[129], unlike fibroblasts which undergo p53-dependent growth arrest[130]. Gene expression profiling also positions ES tumors closer to neural crest stem cells than to bone marrow-derived human mesenchymal stem cells[129]. Yet, it is not known if the neuronal features of ES are derived from the cell of origin or a consequence of EWS-FLI1 expression. In fact, EWS-FLI1 expression in the FN-RMS cell line RD also leads to the upregulation of genes involved in the neural crest development[95].

Unlike neural crest stem cells, mesenchymal cells are normally present in the bone, the most common location of ES[131]. MSCs of the bone marrow can express neural markers spontaneously and can be induced to differentiate towards the neural lineage[131]. Interestingly, knock-down of EWS-FLI1 in ES cell lines switches their expression profile closer to the one of MSCs[132]. Moreover, concomitant silencing of EWS-FLI1 and incubation with differentiation media allows ES cell lines to differentiate into both adipogenic and osteogenic lineages, as do MSCs[132]. Additionally, ectopic expression of EWS-FLI1 in murine MSCs results in cell transformation and in the formation of tumors similar to ES[133,134]. In human MSCs, EWS-FLI1 induces the expression of neuroectodermal markers[135] but fails to develop tumors in mice[136]. Furthermore, expression of EWS-ERG is not well tolerated by mouse progenitor cells[137]. Hence the mystery of ES origin remains unresolved.

Oncogene and transcriptional addiction

Although cancer evolves by the progressive accumulation of mutations, disruption of one gene can completely abrogate the malignant phenotype. Switching on the c-myc oncogene in hematopoietic cells leads to the development of a leukemia highly dependent on it[138]. This observation led to the concept of "oncogene addiction" by Weinstein[139,140] which describes the need for continuous expression of one oncogene for the maintenance of the neoplastic state[141]. Myelocytic leukemias depend on BCR-ABL, melanomas depend on HRAS and the two pediatric sarcomas described above depend on PAX3/7-FOXO1 and EWS-FLI1. The discovery of these "Achilles' heels" of cancer can guide targeted therapy. A good example of this is the use of imatinib to block the activity of BCR-ABL in CML[141]. Nevertheless, many of these "Achilles' heels" are transcription factors (TFs) which are commonly described as undruggable targets due to their nuclear location and lack of enzymatic function[142].

One way of disturbing the action of oncogenic TFs is by targeting transcription itself. Transcription is required for all basic cell processes like survival, growth, and differentiation. However, transcription inhibitors affect more drastically cancer cells than non-transformed cells. Koumenis and GIaccia have shown that RNA polymerase II inhibitors induce growth arrest in non-transformed human fibroblast, while in fibroblasts with deregulated expression of c-MYC they lead to apoptosis[143]. These experiments supported the concept of transcriptional addiction (Fig.15) in cancer cells and made basal transcription machinery a valid target. Indeed, many of the chemotherapeutical drugs used in the clinics are in fact transcriptional inhibitors[144]. The sensitivity of cancer cells to transcription inhibitors can be explained by the fact that anti-apoptotic transcripts have a faster turnover than the pro-apoptotic ones[145]. Moreover, the expression of oncogenes is more easily disturbed by transcriptional inhibition than the one of housekeeping genes[145], due to the epigenetic characteristics of their regulatory regions.

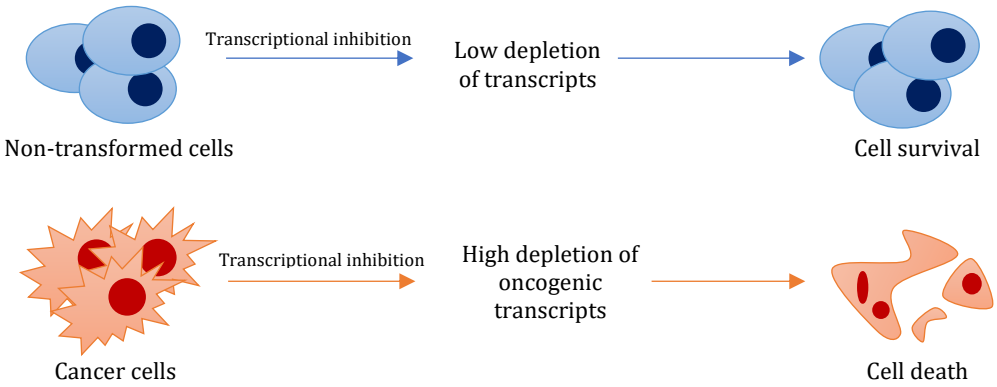

Fig. 15 Differential response of non-transformed cells and cancer cells to transcription inhibition. Adapted from[145].

Epigenetics: the organization of life

The term epigenetics was first applied by C. H. Waddington in 1942[146]. He defined it as "the branch of biology which studies the causal interactions between genes and their products, which bring the phenotype into being"[146].

Epigenetics is the study of heritable changes in gene expression that occur independently of changes in the primary DNA sequence[147]. The epigenetic landscape also shapes the decision-making process during development (Fig.16A)[148] and gives rational to cell differentiation. Even though our cells share the same genotype, we find a broad spectrum of phenotypes whose development is dependent on the epigenome.

1. The epigenetic object: chromatin

Chromatin is composed of DNA and histone proteins organized into units named nucleosomes. These nucleosomes allow the packaging of 2 meters of DNA into a 10-20µm nucleus[149]. The nucleosome is formed by two copies of each histone protein (H2A, H2B, H3, and H4) assembled into an octamer around which 145–147 base pairs (bp) of DNA are wrapped[150]. The H1 histone wraps an additional 20 bp of DNA and forms a 166bp long structure called chromatosome[149]. The nucleosome unit is found every 200 ± 40 bp throughout all eukaryotic genomes, producing an electronic image of beads on a string (Fig.16B). Each nucleosome is separated by an average of 20bp of DNA called the linker DNA[149] and shape the DNA molecule into higher order helices (Fig. 16C).

In conclusion, each chromosome is formed by a string of nucleosomes which are the principal packaging element of DNA and the primary determinant of DNA accessibility.

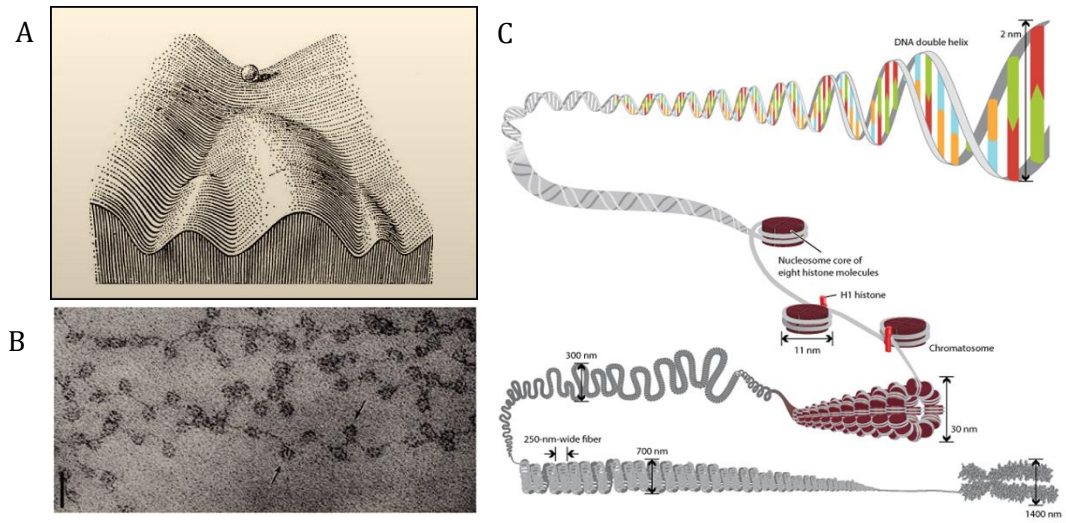

Fig. 16 (A) Model of how the landscape (the epigenome) can define the path (phenotype) of the ball (cell). Adapted from[148]. (B) Beads on a string, an electron micrograph of chromatin. (C) DNA packaging into chromosome. Adapted from[149].

2. Epigenetic regulation of transcription

Transcription, but also DNA-damage repair and replication, requires access to the DNA strands which is blocked by the compact structure of nucleosomes. Hence, there must be mechanisms that allow the transient and reversible opening of chromatin.

Modifications to DNA and histones are dynamically added and removed by chromatin modifying enzymes in a highly regulated manner. These post-translational modifications (PTMs) are then "read" by different chromatin remodeling enzymes that make actively transcribed regions looser and accessible, and repressed regions more closely packed. To date, at least 4 different DNA modifications and 16 classes of histone modifications have been identified[151].

2.1 DNA methylation

Methylation of the 5-carbon residues of cytosine was the first described covalent modification of DNA. DNA methylation represses transcription directly by inhibiting the binding of TFs or indirectly by recruiting proteins associated with repressive chromatin remodeling activities[152]. This covalent modification occurs primarily within centromeres, telomeres, the inactive X-chromosome, and repeated sequences[152]. Interestingly, methylation studies have found many actively transcribed genes to be methylated in their gene bodies, contradicting the dogma that DNA methylation is associated with silencing[151].

The DNA methylation machinery is composed of DNA methyltransferases (DNMTs) which establish and maintain DNA methylation patterns, methyl-CpG binding proteins (MBDs) which "read" the methylation marks[152], and demethylases which erase them[153].

In mammals, DNA is mainly methylated at CpG dinucleotides[154] which are concentrated in CpG-rich DNA stretches named "CpG islands". These islands are then mainly located at the 5'end of genes and occupy 60% of human gene promoters.

In cancer, a global DNA hypomethylation is observed primarily due to the loss of methylation from repetitive regions[152]. However, 5-10% of normally unmethylated CpG islands become abnormally methylated, affecting the expression of genes involved in cell-cycle regulation, tumor cell invasion, DNA repair, chromatin remodeling, and apoptosis[152]. In addition, DNMTs are commonly overexpressed (DNMT1 and DNMT3B) or mutated in cancers (DNMT3A)[151,155,156].

2.2 Histone modifications

Post-translational modifications of histones were first described by Vincent Allfrey in 1964[157]. Since then, many different PTMs have been described, including acetylation, methylation, phosphorylation, ubiquitinylation, sumoylation, ADP ribosylation,

deamination, propionylation, and butyrylation[158]. The most studied modifications are the ones occurring on the N-terminal histone 'tail' regions (Fig.17A), but PTMs in the globular domains of histones have also been described[158].

Histone PTMs influence chromatin compaction and accessibility by changing histone-DNA/histone-histone interactions or by recruiting remodeling complexes and transcription factors(Fig.17B)[158]. Histone acetylation, for instance, unfolds chromatin since it neutralizes the positive charge of lysine, weakening the electrostatic interaction between histones and the negatively charged DNA[159]. Hence, acetylation of histones is associated with an "open" chromatin conformation, which is permissive to transcription[160]. Histone PTMs are also able to directly repress gene expression. Di or trimethylation of H4K20 enhances chromatin condensation, which impairs gene expression[161].

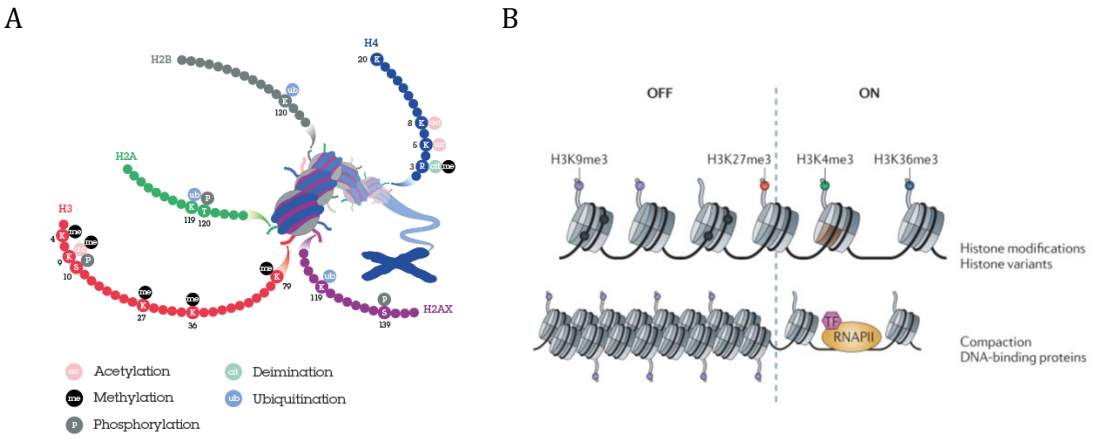

Fig. 17 (A) Representation of histone modifications present in histone tails. Image retrieved from abcam website. (B) Epigenetic regulation of chromatin compaction and transcription. Black dots symbolize DNA methylation. Adapted from[163].

Histone PTMs are constantly being added by enzymes named "writers" and removed by "erasers", which allows a tight control of gene expression and increases cell flexibility and adaptation to environmental changes. Acetylation marks, for example, are added by histone acetyl-transferases (HATs) and erased by histone deacetylases (HDACs). In between, histone modifications are read by specific enzymes with certain "reader" domains. For example, methylation is recognized by chromo, tudor, MBT and PHD domains, and acetylation is recognized by bromodomains[159].

Due to the significance of histone marks in chromatin architecture, Strahl and Allis proposed the existence of a "histone code" where distinct histone modifications can act sequentially or in combination to bring about distinct downstream events which are associated with certain chromatin states[162]. Grossly speaking, there are 2 types of chromatin environments in the genome, silent heterochromatin and active euchromatin (Fig.18)[159]. The histone code suggests that these states are associated with a distinct set of modifications. In mammals, the heterochromatic state is related with low levels of acetylation and high levels of

methylated H3K9, H3K27, and H4K20[159]. Euchromatin, on the other hand, presents a more complex histone code since it is more flexible and changes according to need[159]. Active promoter regions are characterized by sharp peaks of H3K4me3 and histone acetylation, while repressed promoters are decorated with H3K27me3 and H3K9me3[163]. Promoters containing both active (H3K4me3) and repressive marks (H3K27me3), termed "bivalent" or "poised", are usually associated with developmental genes, being inactive in pluripotent cells and induced during differentiation[163]. Enhancers are elements that recruit TFs, RNA Polymerase II (Pol 2) and chromatin regulators to positively influence transcription at distal promoters[163]. These cis-regulatory regions are marked by H3K4me1 and H3K27ac[163]. Finally, actively transcribed genes are distinguished by the presence of H3K36me3 and H3K79me2 in their gene bodies[163].

In cancer, a global loss of H4K16ac, H3K4me3, and H3K20me3, and gains of H3K9me and H3K27me3 have been reported[155]. These alterations have been associated with mutations or overexpression of specific "writers" or "erasers"[151,155]. Hence, many epigenetic drugs have been developed and approved for therapy[164].

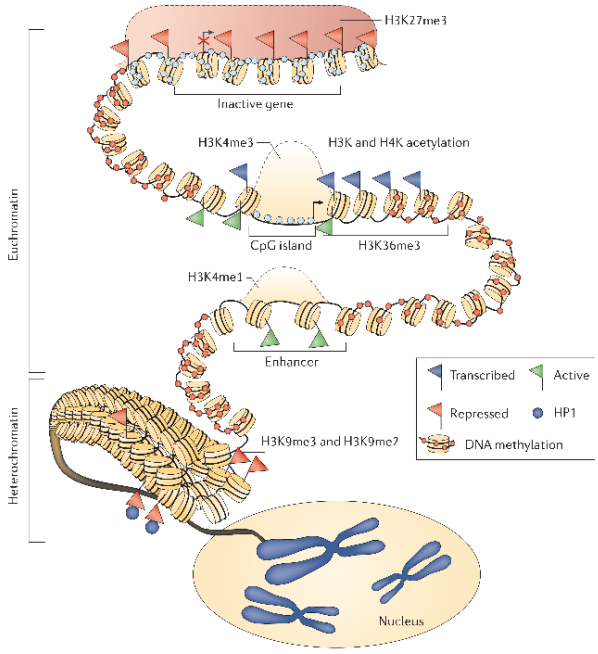

Fig. 18 Epigenetic regulation of chromatin architecture. Adapted from[164].

2.3 Chromatin remodeling

Chromatin remodeling is involved in cell-fate commitment, response to environmental signals, genome replication, and DNA repair[165]. The remodeling process is conducted by helicase-like enzymes that hydrolyze ATP and transform chemical energy into mechanical motion[166]. These enzymes are able to assemble or evict nucleosomes during replication, exchange histone variants, and regulate transcription by changing chromatin accessibility (Fig.19A)[167].

All chromatin-remodeling enzymes belong to the SNF2 family and contain a SNF2 helicase-like ATPase domain[168], which is composed of two RecA-like fold domains (Dexx and HELICc)[169] and is homologous to the ATPase domain of the *Saccharomyces cerevisiae* SNF2 protein[170]. Importantly, SNF2 proteins are not *bona fide* helicases since they lack the ability

to separate nucleic acid strands. Instead, they are DNA translocases that apply torsional strain to DNA and provide the necessary force to remodel nucleosomes[170].

These ATP-dependent remodelers are grouped into 24 subfamilies[166] in accordance with their structure similarities and the presence of other functional domains[171]. The 4 main families (Fig.19B) are the SWI/SNF (characterized by the presence of bromodomains), the imitation SWI or ISWI (containing a SANT domain responsible for histone binding), the chromo and helicase-like domain or CHD (which presents amino-terminal chromodomains that interact with histone tails) and the INO80 (which defining feature is the "split" ATPase domain)[172].

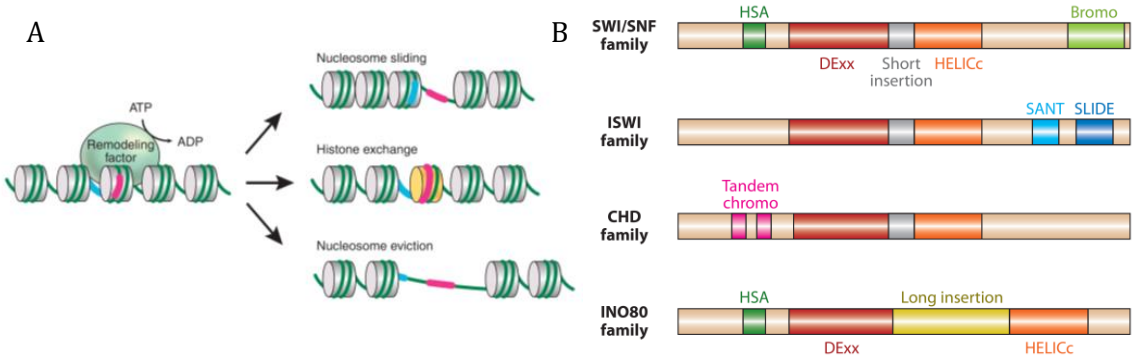

Fig. 19 (A) Chromatin remodeling mechanisms. Adapted from[165]. (B) Structural features of the 4 main subfamilies of the SNF2-like family of ATPases. Adapted from[172].

Despite many ATP-dependent remodelers have been implicated in cancer, the role of chromatin remodeling in tumorigenesis is still not fully understood. BRG1, an ATPase subunit of the SWI/SNF complex, is mutated in cancer cell lines of several tumors types including breast, prostate, lung, pancreas, and colon[173]. Mice with heterozygous mutations of Brg1 are cancer-prone, which suggests a role of Brg1 as a tumor suppressor[173]. Nevertheless, BRG1 can also contribute to cancer development by constraining p53 activity[174]. These findings suggest a context dependent activity for BRG1 in cancer development. More studies are necessary to further investigate the link between chromatin remodeling complexes and tumorigenesis, and to evaluate their therapeutic value.

3. The NuRD complex

The nucleosome remodeling and histone deacetylase complex (NuRD) was first described in 1998 as a multisubunit complex involved in transcription repression[175]. It is highly conserved among eukaryotes, ubiquitously expressed[176], and, to date, the only chromatin remodeler complex bearing both nucleosome remodeling and histone deacetylase activity[177]. The chromatin remodeling function is carried out by CHD3 or CHD4 (chromodomain-helicase DNA binding proteins) while the histone deacetylase function is performed by HDAC1 or HDAC2 (histone deacetylases)[177]. These HDACs share 85% in

sequence identity and are able to remove acetyl groups from lysine residues on histone tails and on other proteins, such as p53[177,178].

Other non-enzymatic subunits of NuRD include MBD2/3 (methyl-CpG-binding domain), RBBP4/7 (retinoblastoma-binding proteins), MTA1/2/3 (metastasis-associated) and GATAD2A/B (GATA Zinc Finger domain containing)[177]. These NuRD members are involved in the assembly and positioning of the complex on the chromatin. MBD2/3 are the smallest members of NuRD[177] and have markedly different functions[179]. While MBD2 binds to methylated DNA[180], MBD3 binds to hydroxymethylated DNA[181]. Furthermore, MBD3 knockout, but not the one of MBD2, leads to embryonic lethality[182]. RBBP4/7 (also known as RbAp48 and RbAp46, respectively) share 92% of sequence homology and direct the complex to histone 3[183]. The MTAs contribute to the assembly of the complex by mediating protein-protein interactions with the HDACs[184,185]. GATAD2A/B interact with the MBDs, increase the repressive function of MBD2, and direct the NuRD to unmodified histone tails[176,177]. Other enzymatic and non-enzymatic subunits have also been associated with this complex such as DOC1-1 (deleted in oral cancer-1) and LSD1 (lysine-specific histone demethylase)[178].

Structural studies suggest that the NuRD is composed by 2 HDACs and MTAs, 4 RBBPs and 1 MBD, GATAD2 and CHD (Fig.20)[176,186]. The ATPase subunits, MTAs and MBDs are mutually exclusive and form complexes with different function[175,178,187]. Some subunits are exclusively present in the NuRD complex (CHDs, MBDs and MTAs), but others take part in other complexes, like HDAC1/2 and RBBP4/7 which are found in the co-repressor complexes SIN3, coREST, PRC2 and CAF-1[178].

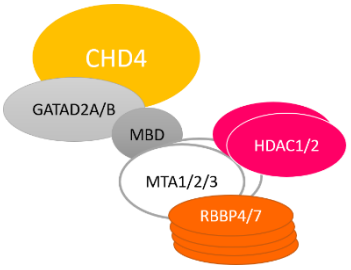

Fig. 20 The NuRD complex.

3.1 The ATPases of NuRD

CHD3 and CHD4 (Mi-2α and Mi-2β, respectively) were first discovered as autoantigens in an autoimmune disease related with cancer development named dermatomyositis[188]. Together with CHD5, which is also found in the context of NuRD, these proteins form the Mi-2 subclass of the SNF2-like ATPase family[189]. In vitro studies demonstrated that CHD4 and CHD3, unlike INO80, ISW1A or CHD1, remodel nucleosomes independently of the linker length and their initial position. Nevertheless, CHD3 and CHD4 remodeling activities result in different nucleosome positioning[187].

Both CHD3 and 4 are ubiquitously expressed and contain a C-terminal ATPase (Fig.21), 2 plant zinc finger homeodomains (PHD1/2), 2 chromodomains and 2 C-terminal domains of unknown function (DUF)[190]. The chromodomains directly bind to DNA while the PHD domains interact with histones[191]. These PHD domains are able to bind the N-tail of H3 with increased affinity when acetylation or methylation of Lys9 is present[192–194]. However, the PHD2 domain of CHD4 fails to bind to H3 in case of methylation of Lys4 or acetylation of Ala1[194]. Importantly, the ATPase activity of CHD4 is dependent on the simultaneous binding of the PHD domains to histones and of the chromodomains to DNA[195,196]. Likewise, the ATPase domain activity promotes the association of the PHD and chromodomains to histones and DNA, respectively[196].

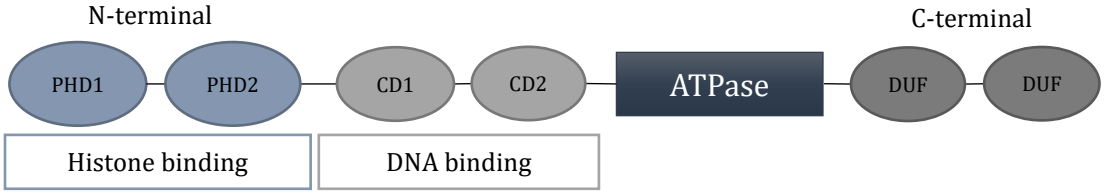

Fig. 21 Schematic representation of the structure of CHD3/4.

Despite their high homology, CHD3/4/5 have different functions. For example, during cortical development, the CHD subunits of NuRD exchange to regulate distinct aspects of differentiation. CHD4 promotes early proliferation of progenitor cells, CHD5 facilitates neuronal migration and CHD3 ensures proper layer specification[197].

3.2 NuRD in DNA-damage repair

To initiate the DNA-damage repair (DDR), the chromatin must be accessible to checkpoint and repair factors. The NuRD is essential to the DDR since it allows an open chromatin conformation at damage sites. In fact, the depletion of CHD4 or MTAs leads to accumulation of spontaneous DNA damage, increased ionizing radiation sensitivity, and impaired homologous recombination repair[198,199]. The remodeling function of CHD4 seems to be crucial for the DDR, since CHD4 mutants with no helicase activity are not able to rescue the depletion phenotype[199].Mechanistically, CHD4 is recruited to the sites of DNA damage by PARP[200] where it participates in the recruitment of BRCA1(Fig.22A). The recruitment of CHD4 is also dependent on the ubiquitin ligase RNF8 and promotes the recruitment of RPA and RAD51[199].

3.3 NuRD in cell cycle progression, DNA replication and genome stability

Besides DDR, the NuRD carries other non-transcription related functions such as regulation of cell cycle, DNA replication, and genome stability.

CHD4/NuRD regulates G1/S transition by controlling p53 acetylation status (Fig.22B)[200]. Indeed, CHD4 depletion leads to hyperacetylation of p53 which increases the expression of p21 and causes G1/S arrest[200].

In lymphocytes, NuRD localizes to pericentromeric heterochromatin (Fig.22C) during S phase together with proteins related to active replication forks such as the proliferating cell nuclear antigen (PCNA) and chromatin assembly protein (CAF1)[201].

The NuRD is also associated with pericentric heterochromatin where it promotes genome stability by inducing nucleosome compaction[202]. This association is driven by the heterochromatin containing HP1 proteins, H3K9me3[203], and sumoylated KAP-1[202]. In fact, loss of CHD4 or MBD3 leads to defects in pericentric heterochromatin assembly and maintenance, which then induces a delayed S-phase progression[204].

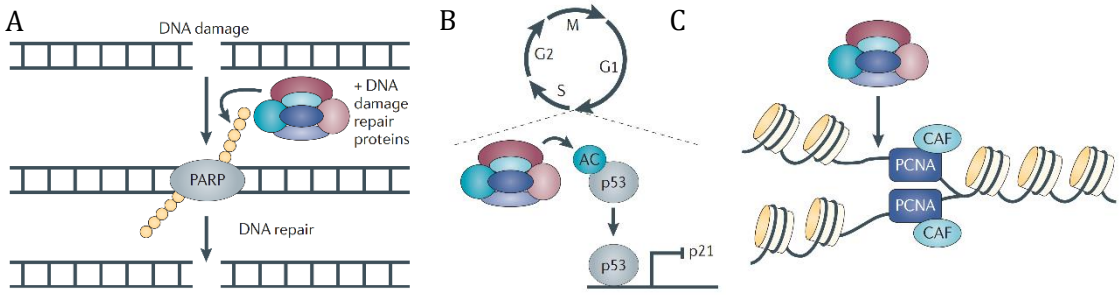

Fig. 22 The NuRD complex functions in DNA replication (A), cell cycle progression (B) and DDR (C). Adapted from[201].

3.4 NuRD in gene regulation and cell fate commitment

When first identified, NuRD was labelled as a repressor complex[175] since MBD2 was found to localize to methylated promoters from where it induced gene silencing through collaboration with histone deacetylases[205]. However, more recently, many studies have shown the potential of NuRD as an activator. Sequencing studies position the MBD3 component at active promoters and H3K27ac rich enhancers[206]. Moreover, during hematological development, NuRD mediates both activating and repressive functions. For example, in erythroid cell lines, NuRD coregulates FOG-1 dependent transcriptional activation[207], while, during T-helper 2 cell differentiation, Chd4 forms a complex with Gata3 and p300 that mediates both positive and negative gene regulation[208].

The NuRD complex plays many roles in cell differentiation. In embryonic stem cells (ESCs), NuRD-mediated deacetylation is associated with PRC2 recruitment and subsequent methylation of H3K27, which determines the transcription status of poised genes[209]. During muscle development, CHD4 is required for the correct differentiation of cardiac muscle. In fact, Chd4 depletion leads to aberrant expression of the skeletal muscle program and causes cardiomyopathy in mice[210]. Moreover, CHD4-mediated repression of skeletal and smooth

muscle myofibril isoforms is required for normal cardiac sarcomere formation[211]. In neural progenitor cells (NPCs), Chd4 interacts with Ezh2 and silences glial markers, preventing astroglial differentiation[212]. Moreover, NuRD interacts with the TF LHX2 and controls the expression of Sox11 and Fezf2, two TFs essential for cortical neuron development[213].
Interestingly, some NuRD elements seem to have contradictory functions in development. Mbd3, in the absence of self-renewal signals, silences the expression of pluripotency genes in ESCs and allows the cell to respond to differentiation[214]. On the other hand, Chd4 was reported to be essential for the self-renewal of ESC[215].

3.5 NuRD and cancer

Many NuRD components are overexpressed in cancer and involved in tumorigenesis. Less frequently, NuRD components are mutated during the tumorigenesis process. Nonetheless, 19% of uterine serous carcinomas and 17% of endometrial cancers carry CHD4 loss-of-function mutations[216].
MTA1 overexpression is associated with aggressiveness of breast, liver, colon, pancreas, and prostate cancers[217]. In breast cancer, MTA1 and MTA3 carry opposing functions. MTA1 facilitates tumor progression through interaction and repression of the estrogen receptor transcriptional function, while MTA3 acts as a tumor suppressor by interfering with the expression of SNAI, a crucial TF that promotes epithelial-mesenchymal transition (EMT). Furthermore, MTA1 expression increases during breast cancer progression, while the one of MTA3 decreases[201]. In gastroesophageal junction adenocarcinoma, low levels of MTA3 are also associated with a poorer prognosis[218]. MTA1 is a downstream target of the oncoprotein c-MYC and is essential for c-MYC driven cell transformation[219]. These findings point to both an oncogenic and tumor suppressor function of NuRD members. In fact, on one hand, MBD2 contributes to the silencing of hypermethylated promoters of tumor suppressor genes, including the CDKN2A locus in colon cancer[220], and, in breast cancer, the tumor suppressor SALL1 interacts with MBD3 to inhibit tumorigenesis and metastasis[221]. Thus, the role of NuRD in cancer seems to be subunit and context dependent and requires further investigation.

4. Targeting epigenetic regulators

Since epigenetic regulators control genome accessibility in a reversible manner, inhibitors targeting these regulators are promising anticancer strategies. In fact, many small molecules targeting DNA methyltransferases, histone methylases, demethylases and deacetylases have been developed and some are already in use in the clinics.

Two DNMTs inhibitors, azacitidin and decitabine, were approved by the FDA and are used in myelodysplastic syndrome (MDS) therapy. Nonetheless, there is still room for improvement considering that these drugs lack oral bioavailability, have a short half-life and lack specificity as they bind irreversibly to all catalytically active DNMTs[222]. HDAC inhibitors are also used in the clinics. Romidepsin and vorinostat have been approved for the treatment of cutaneous T-cell lymphoma. Nevertheless, the mechanism of action of HDAC inhibitors is still debatable. Their anti-tumor activity might be due to alterations in transcription or via regulation of the acetylation status of non-histone proteins, like p53[223]. Although, not in the clinics yet, targeting of the histone methyltransferases (HMTs) EZH2 and DOT1L holds great promise in cancer therapy. EZH2 is the catalytic component of the Polycomb protein complex (PRC2) and is responsible for the H3K27me3 mark. In prostate, breast, kidney, and lung cancer, increased levels of EZH2 are associated with poor outcome[224]. DOT1L specifically methylates H3K79 and contributes to the transcriptional activation of HOXA10 and MEIS1, which are required for leukemia initiation[225]. Emerging preclinical data has shown the therapeutic potential of the pharmaceutical inhibition of the histone demethylase LSD1(lysine-specific histone demethylase 1A) in AML. There, LSD1 is required to sustain the expression of the target genes of the oncoprotein MLL-AF9. Besides AML, this epigenetic effector is highly expressed in prostate cancer, undifferentiated neuroblastoma, estrogen-negative breast cancer, bladder cancer and colorectal cancer[224]. LSD1 inhibitors are expected to go soon into phase I clinical trials.

4.1 BET-bromodomain inhibition

Bromodomains are a small family of proteins that recognize and bind to acetylated lysine residues on histone tails. There, they act as scaffold of large multi-component complexes that regulate chromatin accessibility and transcription[224].

The demonstration of the druggability of the reader bromodomain and extra-terminal (BET) subfamily (BRD2/3/4/T) has intensified the interest in these epigenetic regulators as therapeutic targets[226]. BET-bromodomain inhibitors cause dramatic antiproliferative and pro-apoptotic effects in a variety of cancer cell lines and tumor growth delays in mouse xenograft models. Especially susceptible to these inhibitors are hematological malignancies including AML, Burkitt's lymphoma, multiple myeloma, and B-cell ALL[155,226], where these inhibitors directly silence MYC expression. Interestingly, in neuroblastoma, these inhibitors affect MYCN expression and cause tumor cell death[227,228]. Consequently, two BET-bromodomain inhibitors, GSK 525761 and OXT015, are now in phase I trials for carcinoma and hematological malignancies, respectively[229].

Aim

Gene expression and its spaciotemporal regulation are essential for the correct development of cells. During tumorigenesis, aberrant transcription factors or dysregulated signaling pathways change gene expression programs and cause cell transformation. Tumor cells become addicted to the expression of certain oncogenes and to a high level of transcription, which also makes them vulnerable to transcription inhibition and creates therapeutic opportunities. This is particularly relevant for the two pediatric sarcomas studied here, fusion-positive rhabdomyosarcoma and Ewing sarcoma.

FP-RMS and ES are driven by a fusion TF (PAX3-FOXO1 and EWS-FLI1, respectively) that binds to cis-regulatory regions in the genome, mainly enhancers, and, from there, controls gene expression. Targeted therapies developed for these sarcomas have been focused on the fusion protein interactors or on the inhibition of target proteins. These approaches carry many limitations, are prone to resistance development, and have been so far unsuccessful.

A growing number of epigenetics studies has shown that the gene expression programs regulated by aberrant TFs are highly dependent on epigenetic regulators. Hence, targeting these regulators represents a valid strategy to interfere with the oncogenic signature. The development of JQ1, a small-molecule BRD4 inhibitor, provided evidence of the feasibility of such approach. Since its development, many in vitro and in vivo studies demonstrated that a broad number of cancers are sensitive to BET-bromodomain inhibition.

We and others identified the chromatin remodeler CHD4 (chromodomain-helicase-DNA-binding protein 4) as a new epigenetic coregulator of PAX3-FOXO1 and EWS-FLI1. Chromatin remodelers are ATPases capable of using energy drawn from ATP hydrolysis to move nucleosomes along the DNA. Consequently, they control chromatin accessibility to the transcription machinery and are essential to gene expression control. Despite the relevance of chromatin remodeling in transcription, it remains a poorly studied branch of epigenetics and is often overlooked by the drug discovery field.

Here, we aim to explore chromatin remodeling and changes in DNA accessibility as valid strategies for cancer therapy. To do so, we aim to investigate the mechanisms employed by CHD4 in the control of the oncogenic signatures of both FP-RMS and ES. Through chromatin immunoprecipitation assays followed by genome-wide sequencing, we describe in detail the epigenetic mechanisms used by PAX3-FOXO1 to control gene expression in FP-RMS. In addition, by performing DNase hypersensitivity assays, we study the role of CHD4's nucleosome remodeling activity in PAX3-FOXO1-driven transcription. We explore also the potential of CHD4 as therapeutic target in both FP-RMS and ES. At last, we aim to develop

the first CHD4 small molecule inhibitor, using structure-based drug design, to validate CHD4 as druggable new therapeutic target in cancer.

With this study, we hope to highlight the role of chromatin remodeling in oncogenic gene expression, to establish CHD4 as an unexpectedly broad tumor susceptibility and to boost chromatin remodeling as a new field of potential cancer therapeutic targets.

Chapter Two: Manuscripts

Manuscript 1

The chromatin remodeler CHD4 regulates super-enhancer driven gene expression and cooperates with BRD4 to drive fusion positive pediatric sarcomas

Joana G. Marques[1], Berkley E. Gryder[2], Blaz Pavlovic[1], Yeonjoo Chung[1], Quy Ngo[1], Young Song[2], Katharina Holste[1], Dominik Laubscher[1], Marco Wachtel[1], Javed Khan[2], Beat W. Schäfer[1]

[1] Department of Oncology and Children's Research Center, University Children's Hospital Zurich, Zurich, Switzerland
[2] Genetics Branch, National Cancer Institute, National Institutes of Health, Bethesda, MD, USA

Conflict of interest: The authors declare no conflict of interest.

Manuscript ready for submission.

Corresponding Author: Beat W. Schäfer

Department of Oncology, Children's Hospital Zurich
Steinwiesstrasse 75
8032 Zurich, Switzerland
Beat.schaefer@kispi.uzh.ch
Phone: +41 (44) 266 7553
Fax: +41 (44) 634 8859

Author contributions: JGM, BP, YC, YS, KH, DL acquired data. JGM, BEG and QN performed bioinformatic analysis of sequencing data. JGM, BEG and MW contributed to the analysis and interpretation of data. JGM, BEG, MW, JK and BWS made substantial contributions to the conception, design and intellectual content of the paper. JGM and BWS wrote the paper.

ABSTRACT

In cancer, aberrant transcription factors and/or dysregulated signaling pathways lead to changes in gene expression that drive tumorigenesis. Hence, targeting transcription through inhibition of epigenetic regulators is a valuable cancer therapeutic strategy, as exemplified by the recent success of BET-bromodomain inhibition.

Here, we explored the potential of chromatin remodelers as therapeutic targets, since they have been overlooked so far as drug targets. As models of transcriptional deregulation, we interrogated two pediatric sarcomas, fusion-positive rhabdomyosarcoma (FP-RMS) and Ewing sarcoma (ES), which are driven by oncogenic transcription factors, PAX3-FOXO1 and EWS-FLI1 respectively. In both tumors, we characterized the epigenetic role of the chromatin remodeler CHD4, which was previously identified as a necessary cofactor for PAX3-FOXO1and EWS-FLI1 activity. We demonstrate that CHD4 co-localizes with PAX3-FOXO1 at super-enhancers together with BRD4 and the NuRD members HDAC2 and RBBP4. Surprisingly, CHD4 directly interacts with BRD4 at these sites and generates a chromatin architecture permissive for the binding of the fusion protein and of its cofactor BRD4. This allows expression of the oncogenic program of PAX3-FOXO1 and permits tumor cell survival. The other members of the NuRD complex influence transcription to minor extend and have a less relevant role in FP-RMS cell viability. Our results suggest that CHD4 depletion can disrupt super-enhancer driven gene regulation by interfering with chromatin architecture in a NuRD-independent manner. Other studies have shown that CHD4 interacts and coregulates EWS-FLI1 target genes and here, we show that depletion of this remodeler induces ES cell death in vitro and impedes tumor growth in vivo. Finally, analysis of genome-wide cancer dependency databases identifies CHD4 as general novel cancer vulnerability.

Together, our findings highlight chromatin remodeling as possible novel means to influence super-enhancer driven oncogenic gene expression and to impair tumor growth.

Significance

- CHD4 localizes to super-enhancers and regulates their influence on gene expression
- BRD4 interacts and co-localizes with CHD4 at super-enhancers
- CHD4 acts independently of NuRD to regulate PAX3-FOXO1 target genes
- Silencing of CHD4 impairs Ewing sarcoma tumor growth in vivo
- A vast number of cancer cells are susceptible to CHD4 inhibition

Keywords:

CHD4, chromatin remodeling, transcriptional regulation, super-enhancers, BRD4

INTRODUCTION

Transcriptional programs control cell identity and survival. In cancer, oncogenic transcription factors (TFs) deregulate gene expression to cause cell transformation. As a result, cancer cells become addicted to not only the continuous expression of the oncogenic TF but also to high levels of transcription[1,2]. Indeed, targeting transcription represents a valid therapeutic strategy since cancer cells are more susceptible to transcription inhibition than healthy cells[3] even though transcription regulators are broadly expressed and play essential roles in healthy cells.

Epigenetics provides a new field of possible drug targets that can interfere with transcription. In fact, epigenetic regulators control the selective accessibility of TFs and of the transcriptional machinery to specific genomic elements, such as enhancers and promoters. These regulators are enzymes that can alter the chromatin chemically (via DNA methylation and post-translational modification of histones) or structurally (via chromatin remodeling)[4]. Since many epigenetic effectors are associated with tumorigenesis through mutation or overexpression, new classes of drugs have been developed[5]. DNA methyltransferase (DNMT) inhibitors are used as first-line treatment of myelodysplastic syndrome[6] and histone deacetylase (HDAC) inhibitors are currently used in T-cell lymphoma therapy[7,8]. More recently, pharmacological inhibition of bromodomain-containing protein 4 (BRD4), a BET-bromodomain protein that acts in concert with the positive elongation factor complex (P-TEFb) and is involved in the control of RNA Pol2 pause release[9,10], was shown to drastically impact oncogenic transcription, rendering many tumors susceptible to its inhibition[11–14].

Within epigenetic effectors, chromatin remodelers have been so far not considered as possible drug targets for cancer therapy. Chromatin remodelers regulate gene expression by controlling the packaging and unpackaging of DNA in the nucleosomes at regulatory sites[15]. Usually, chromatin remodelers work as protein complexes that contain SNF2-like ATPases. These enzymes use chemical energy drawn from ATP hydrolysis to cause mechanical movement of nucleosomes, and, hence, control DNA accessibility of the transcriptional machinery[16,17]. Therefore, also chromatin remodeling enzymes might constitute valuable targets to influence oncogenic transcription.

Fusion-positive rhabdomyosarcoma (FP-RMS) and Ewing sarcoma (ES) are two pediatric tumors that carry a low mutational burden and are transcriptionally addicted to the oncogenic program of the fusion TFs PAX3-FOXO1 and EWS-FLI1, respectively[18–23]. This makes these sarcomas susceptible to epigenetic cofactors depletion. Indeed, BRD4 inhibition was shown to affect the expression signature of both fusion proteins and, consequently, to impair proliferation of both tumors[24–26].

Recently, we have identified the chromatin remodeler CHD4 (chromodomain-helicase-DNA-binding protein 4) as a crucial coregulator of PAX3-FOXO1 target genes which is essential for FP-RMS cell viability[27]. However, the exact mechanisms by which CHD4 influences gene expression in this tumor have not been clarified. Similarly, in Ewing sarcoma, CHD4 was shown to help the repressive function of EWS-FLI1, but no mechanistic details were provided[28].

In this study, we hypothesized that chromatin remodeling might constitute a novel valid strategy to interfere with oncogenic transcription and hence tumor development. To address this, we focused on the mechanisms by which CHD4 influences gene expression in FP-RMS and ES and describe a new role of CHD4 in the regulation of super-enhancer-driven gene expression. CHD4 co-localizes with PAX3-FOXO1 at super-enhancers, where it regulates DNA accessibility and, therefore, binding of the fusion protein and of its cofactor BRD4 to the chromatin. Moreover, we demonstrate the potential of CHD4 inhibition in Ewing sarcoma and highlight this chromatin remodeler as a promising new tumor target in a wide variety of cancer types.

RESULTS

CHD4 localizes at enhancers together with PAX3-FOXO1 and other members of the NuRD complex

To localize the chromatin remodeler CHD4 in the FP-RMS genome, we performed chromatin immunoprecipitation assays followed by DNA sequencing (ChIP-seq) in the FP-RMS cell line RH4. Since all CHD4 antibodies currently available lack specificity, we introduced a 3xflag peptide in-frame at the N-terminal of the endogenous CHD4 protein via CRISPR/Cas9 mediated repair. Immunoblotting confirmed the insertion of the flag peptide (Fig.S1A) while immunofluorescence showed that flagged CHD4 (N-CHD4) was nuclear (Fig.S1B). In addition, we confirmed that the flag insertion did not affect cell proliferation (Fig.S1C). This allowed us to perform ChIP-seq experiments with an anti-flag antibody to track CHD4 location. To localize PAX3-FOXO1 (P3F) in the genome we used a monoclonal breakpoint specific antibody[29]. Furthermore, we performed ChIP-seq for two members of nucleosome remodeling and deacetylase complex (NuRD), HDAC2 and RBBP4[30], as CHD4 is commonly associated with this complex. Since BRD4 co-localizes and regulates P3F target genes, we used previously published ChIP-seq data for this epigenetic factor to investigate its relation to CHD4[24]. Then, using ChIP-seq data from a comprehensive set of histone marks, we defined 16 chromatin states in FP-RMS[31] (Fig.1A). The overlay of the ChIP signals of CHD4, RBBP4, HDAC2, BRD4, and P3F to these chromatin states demonstrated that CHD4, similarly to P3F and BRD4, localized mainly to enhancers and less so to promoters. RBBP4 and

HDAC2, were also present in enhancers but, unlike CHD4, showed an additional strong prevalence for promoters. Besides enhancers, our data revealed a strong presence of P3F in the transcription transition state. The overlap between the ChIP peaks of P3F, NuRD (represented by the co-localization of HDAC2 and RBBP4) and CHD4 (Fig.1B) showed that 41% of P3F peaks (1531 out of 3694 of the peaks called for P3F) co-localized with CHD4 and NuRD. An example of the co-localization of CHD4, RBBP4, HDAC2 with P3F and BRD4 at an intronic enhancer of the P3F target gene ALK is displayed in Figure S2A.

To further characterize the co-localization of CHD4 with P3F, we divided the genome in P3F-only, CHD4-only and CHD4+P3F locations (Fig.1C/D). Interestingly, CHD4 co-localized with P3F in roughly half of all P3F binding sites (2011 out of 3694 P3F peaks). CHD4+P3F regions were rich in the enhancer marks H3K4me1, H3K27ac and H3K9ac, and poor in the repressive mark H3K27me3 and the promoter mark H3K4me3, confirming that these regions are indeed active enhancers. Moreover, CHD4+P3F binding sites were located distal to the transcription start sites (TSSs) and were mainly found in intergenic and intronic locations (Fig.S2B/C). As expected for activel regions, CHD4+P3F locations are more sensitive to DNase digestion than P3F- or CHD4-only regions (Fig.1C), suggesting a more open chromatin conformation. The same locations are also densely occupied by BRD4, MYCN and the master transcription factors (MTFs) of myogenesis MYOG and MYOD1 (Fig.1D). Furthermore, motif analysis (Fig.1E) revealed that CHD4+P3F locations were enriched in motifs related to myogenic transcription factors (MYF5, MYOD, and MYOG). This is consistent with the myogenic features usually ascribed to this tumor[19]. Finally, GREAT ontology analysis (Fig.1F) on CHD4+P3F locations revealed that these regions are related to BMP signaling, neural tube closure and formation, and neuronal differentiation, which nicely reflects the role of PAX3 during normal neural development[32]. In addition, these findings suggest a function of CHD4 during development as well as in tumorigenesis.

In conclusion, CHD4 was found to co-localize at enhancers with P3F, BRD4 and the NuRD members, HDAC2 and RBBP4, that displayed an open chromatin structure as well as at locations with more repressive characteristics in CHD4-only regions. These results indicate that the function of this remodeler is likely context dependent.

CHD4 binds to super-enhancers and interacts with BRD4

Super-enhancers (SEs) are enhancer clusters that are abundantly populated by TFs and cofactors like the mediator complex, the acetyltransferase p300, BRD4, and the cohesion complex[33]. During normal development, SEs regulate cell identity, while in cancer, they drive high levels of expression of genes related to cell proliferation and transformation[33,34]. Since P3F regulates the expression of oncogenes by binding to SEs together with other MTFs

of myogenesis and BRD4[24], we evaluated the presence of CHD4 at SEs. Strikingly, CHD4 was found in 600 out of the total 810 SEs described in FP-RMS, and co-localized with P3F in 408 out of the 452 SEs where the fusion was present (Fig.1G). HDAC2 and RBBP4 were present in a total of 805 SEs and co-localized with BRD4 in a total of 770 SEs (Fig.1G/H). This is exemplified by the MYOD1 SE in which P3F, RBBP4, HDAC2, BRD4, and CHD4 co-occupy the same genomic location (Fig. S2D).

Recently, a direct physical interaction between CHD4 and the extra-terminal (ET) domain, a highly conserved domain, of BRD3 was reported[35]. Since BRD4 and CHD4 co-localized in FP-RMS at enhancers and SEs, we hypothesized that CHD4 might co-regulate P3F target genes by direct interaction with BRD4. To test this hypothesis, we performed co-immunoprecipitation assays using N-terminal flag tagged endogenous BRD4, and N- and C-terminal flag tagged endogenous CHD4. The knockins of the 3xflag sequence were obtained by CRISPR/Cas9 mediated repair and were validated as shown in Fig.S1 and S3. Immunoprecipitation of N-BRD4 (Fig.1I/J) clearly coprecipitated CHD4, as well as the NuRD members HDAC2 and MTA2. In the reverse setting, CHD4 immunoprecipitation was able to pull-down BRD4 (Fig.1K) as well as MTA2 and HDAC2, as expected. Together, these findings suggest that CHD4 might influence chromatin architecture to regulate enhancer-driven gene expression by collaborating with BRD4.

CHD4 allows an open chromatin conformation at super-enhancers which is permissive to active transcription

CHD4 is a nucleosome remodeler capable of moving nucleosomes along the DNA[36]. Thus, we hypothesized that CHD4 regulates transcription by changing the chromatin architecture at regulatory locations. To investigate this in detail, we performed DNase hypersensitivity (DNase HS) assays upon CHD4 or P3F depletion using RH4 cells with doxycycline-inducible shRNAs targeting either CHD4 or P3F, that were previously described [27]. CHD4+P3F regions lost DNase HS signal after depletion of either CHD4 or P3F whereas CHD4-only locations were generally closed and therefore not affected by the silencing. As expected, P3F-only locations lost DNase HS signal after P3F silencing and were not as affected by CHD4 depletion. Strikingly, when analyzing SE regions only, we observed that these were drastically affected by CHD4 but not by P3F deletion (Fig.2B), suggesting that occupancy by CHD4 is necessary for P3F binding at SEs.

Hence, we next investigated whether changes in chromatin conformation would alter the genomic localization of P3F. To accomplish this, we performed ChIP-seq with a P3F breakpoint-specific antibody[29] upon depletion of CHD4 or P3F. Indeed, while CHD4 depletion caused only small changes in the P3F occupancy when looking at all CHD4+P3F

regions (Fig.2C), it showed clear reduction in P3F occupied sites when analyzing SEs separately (Fig.2D). As expected, P3F depletion reduced its occupancy in all regions. ChIP and DNase signals in the ALK and MYOD1 SEs, illustrating the displacement of P3F from SE loci upon CHD4 depletion, are shown in Figures 2E and F.

To evaluate if changes in chromatin accessibility at SEs cause displacement of epigenetic cofactors, we selected regions where NuRD proteins and BRD4 bind together with CHD4 and P3F and performed ChIP-qPCR assays after silencing of CHD4 or P3F. We observed that CHD4 and P3F depletion impaired the binding of both BRD4 and HDAC2 to the promoter of CDH3 and to the SE present in the intronic region of ALK but not to the one in ASS1.

These findings suggest that the remodeling function of CHD4 is essential to keep the chromatin open at super-enhancers which allows the positioning of the oncogenic driver PAX3-FOXO1 and of its cofactor BRD4 and the expression of the fusion gene signature.

NuRD-only localizes to promoters and participates in cellular homeostasis

Recently, the role of CHD4 as a core member of NuRD has been challenged[37]. Since we also observed that CHD4, HDAC2, and RBBP4 did not co-occupy all genomic locations (Fig.1B), we aimed next to investigate the characteristics of CHD4-containing NuRD and NuRD-only regions (Fig.S4A). We observed that NuRD-only regions were predominantly found in promoters, as indicated by the strong signal of the promoter mark H3K4me3 (Fig.S4A and B). Indeed, NuRD-only peaks localized closer to TSSs than CHD4+NuRD locations (Fig.S4C). Interestingly, BRD4 co-localized with NuRD at both enhancer and promoter locations (Fig.S4A), suggesting that BRD4 might also be able to interact with members of the NuRD in the absence of CHD4. As an example, the TOP2A promoter occupied by NuRD and BRD4, but not by CHD4 is shown in Figure S4D. Gene ontology analysis revealed that the promoters occupied by NuRD-only regulate processes related to cell homeostasis such as RNA processing and ribosome biogenesis (Fig.S4E).

These data suggest that NuRD locates differently depending on the presence of CHD4. In the presence of CHD4, NuRD localizes to enhancers and in the absence of it to promoters. Since CHD4 and CHD3 are mutually exclusive subunits of the complex, the promoter regions enriched in HDAC2 and RBBP4 might represent CHD3-containing NuRD locations. In fact, studies have shown that CHD3 and CHD4 have distinct nuclear localization and also carry different remodeling activities[38]. At promoters, NuRD, possibly with CHD3, regulates housekeeping gene expression and at enhancers it might drive P3F oncogenic signature together with CHD4.

CHD4, but not NuRD, is essential for FP-RMS cell viability

Previously, we have demonstrated that silencing of CHD4 leads to FP-RMS cell death and tumor regression in vivo[27]. Here, we discovered that CHD4 co-localizes with P3F at SEs together with two members of the NuRD complex, HDAC2 and RBBP4. This raised questions regarding the role of NuRD in P3F-driven gene expression and FP-RMS viability. Therefore, we probed all known NuRD subunits for their role in FP-RMS cell proliferation. We established a NuRD-centered CRISPR/Cas9 screen in which we used FP-RMS cells stably expressing Cas9 (Fig.3A) and transduced them either with five RFP-labelled sgRNAs per NuRD member or with a BFP-labelled control guide (sgAAVS1). Two days after transduction, the blue and red population were mixed 1:1, and after 12 days, the propagation of the populations was assessed by FACS analysis. Guide RNAs targeting BRD4 were used as positive control[24]. The results of this screen are depicted in Figure 3B as the ratio between red and blue population at day 12 normalized to day 2. We considered that the knockout had an effect in cell proliferation when at least 3 sgRNAs caused a ratio reduction superior to 25%. According to these criteria, we observed that proliferation of FP-RMS was reduced not only by targeting CHD4 but also by HDAC1, RBBP4/7, and GATAD2A/B. Importantly, none of the MBD or MTA proteins, which are unique members of the NuRD[39] complex, altered FP-RMS cell proliferation (Fig.3B).

Since RBBP4 knockout showed comparable effects on proliferation to CHD4's, we studied it further by establishing FP-RMS cell lines containing two doxycycline-inducible lentiviral shRNAs targeting RBBP4. Silencing was confirmed on protein and mRNA levels 72 hours after doxycycline induction (Fig.S5A/B). As expected, RBBP4 depletion decreased FP-RMS cell proliferation, as assessed by crystal violet and WST1 assays (Fig.S5C). Nonetheless, unlike CHD4 and P3F depletion, RBBP4 knock-down did not induce FP-RMS cell death (Fig.3C/D, S5D) or influenced the expression of 5 out of 6 selected P3F target genes (Fig.3E-G). Hence, reduction of proliferation triggered by RBBP4 depletion occurs independently of the P3F signature. Taken together, these observations suggest that the loss of cell viability induced by CHD4 depletion does not dependent on the formation of the NuRD complex.

CHD4 regulates the oncogenic program of FP-RMS

To further investigate the transcriptional dependency on CHD4 in the context of the fusion gene signature, we performed RNA-seq experiments upon CHD4 or P3F depletion. The validation of the knockdowns obtained at 48hrs, both on protein and mRNA levels, is presented in Figure 4A and B. We observed that CHD4 does not regulate the expression of P3F itself, hence results obtained after CHD4 knockdown are not indirectly caused by reduction of the fusion protein.

Unsupervised clustering analysis (Pearson correlation) of the genes differentially regulated by CHD4 (fold change of at least 25% and FDR of 1%) identified a similar number of genes upregulated (3080) and downregulated (3220) upon CHD4 silencing with a high degree of reproducibility (Figure 4C). These results suggested both an activating as well as repressing function of CHD4. Of all CHD4 regulated genes, 2388 were coregulated by P3F (1238 downregulated and 1150 upregulated).

Gene set enrichment analysis (GSEA) performed on the CHD4+P3F transcriptome revealed an enrichment of the signatures of P3F-controlled enhancers and of FP-RMS (Fig.4D), which is consistent with the role described for CHD4 as cofactor of P3F in the regulation SE-driven expression. When we considered all genes differentially regulated by CHD4, we found an enrichment of MEF2 and embryonic stem cell signatures (Fig.4E). MEF2 comprises a family of transcription factors that collaborates with other myogenic transcription factors, such as MYOD, to drive skeletal muscle differentiation[40]. Therefore, these results supported a role of CHD4 in normal skeletal muscle differentiation[41]. In fact, during muscle differentiation, this chromatin remodeler, through binding to distal regulatory regions, directly regulates the expression of genes required for cardiac muscle development[42].

Unexpectedly, amongst the genes regulated by CHD4 we identified also the gene expression signatures of EWS-FLI1 and Ewing sarcoma (Fig.4E). This finding suggests some commonality amongst the oncogenic signatures of these sarcomas and could potentially help to identify physiologically relevant target genes.

CHD4 depletion drastically decreases Ewing sarcoma cell viability and reduces tumor growth in vivo

CHD4 has been shown to collaborate with EWS-FLI1 to repress some of its target genes[28]. Moreover, this remodeler is not only highly expressed in RMS but also in ES (Fig.5A) and regulates some ES genes in FP-RMS. Thus, we explored its possible role for ES cell viability. To this end, we established two ES cell lines, SKNMC and A673, to stably express two doxycycline-inducible shRNAs targeting CHD4. Validation of CHD4 depletion through doxycycline incubation is shown in Figure S6A and B. Indeed, and comparable to FP-RMS, CHD4 depletion deeply impaired ES cell proliferation, as shown by crystal violet and WST1 assays, in both ES cell lines and with both shRNAs (Figure S6C-F), and induced tumor cell death (Figure S6G).

To assess the effect of CHD4 depletion on tumor growth in vivo, NOD/SCID mice were engrafted subcutaneously with SKNMC cells containing either the scramble control shRNA construct or the shCHD4#1 (Fig.5B). Then, mice, with established tumors of an average size of 60mm^3 per group, were treated intraperitoneally with a doxycycline solution (53.3

mg/kg) for two consecutive days, and, from then on, fed with doxycycline-supplemented food. Control mice were injected intraperitoneally with vehicle (PBS solution) and continued on a doxycycline-free diet. Immunohistochemistry assays validated the silencing of CHD4 in vivo (Fig.5C). CHD4 depletion clearly caused a delay in tumor growth (Fig.5D) and increased overall survival (Fig.5E). Interestingly, the doxycycline-treated tumors bearing the shCHD4#1 construct that grew back to 1000mm^3 re-expressed CHD4 up to control levels (Fig.5F). Thus, CHD4 re-expression was the likely cause of the outgrowth of these tumors. These results validate CHD4 as a possible therapeutic target also in ES.

Large-scale genome screens suggest CHD4 as a broad tumor dependency

CHD4 has been implicated in the viability of different tumors. Here, we described its role in Ewing sarcoma cell survival and in the regulation of P3F-driven gene expression in FP-RMS. Furthermore, we observed that CHD4 collaborates with BRD4 for SE-driven gene expression regulation. Therefore, we questioned if CHD4 constitutes a general tumor susceptibility as demonstrated for BRD4. To answer this, we analyzed data generated by two genome-wide screens which aimed to identify cancer susceptibilities, the DRIVE[43] and the Achilles projects (https://portals.broadinstitute.org/achilles). In both projects, thousands of targets were screened using a shRNA- (DRIVE) or CRISPR-based (Achilles) library in nearly 400 cancer lines (Supplementary Table 1 and 2). Then, dependency scores were calculated and made publicly available.

First, we analyzed the data available for all NuRD complex members in both datasets (Fig.6A and B). We observed that CHD4 impairs tumor cell viability in the majority of the cancer cell lines used in both projects (in the DRIVE database 323 out of 387 cancer cell lines had a score lower than -3 upon CHD4 silencing; in the Achilles database 386 out of 391 cancer cell lines tested had a score lower than -0.5 upon CHD4 knockout). Other NuRD members showed either no effect in all tumor cells or selected tumor dependencies, apart from RBBP4 and GATAD2A depletion which had a broad effect on cancer cell viability in the Achilles platform only (Fig.6A).

Next, we analyzed the effect of CHD4 with the one of all other SNF2-like chromatin remodelers and BRD4[17]. In both databases, only INO80 and BRD4 depletion showed a similar general tumor dependency as the one observed for CHD4 (Fig.6C and D).

This analysis suggests a vast tumor susceptibility to CHD4 depletion, beyond the two pediatric sarcomas here studied, and suggest it to be a promising new tumor target in a wide variety of cancer types

DISCUSSION

Cancer is a very heterogenous disease with a high degree of variability within tumor types, which makes it difficult to identify common drug targets. Recently, attention has been brought to epigenetic targeting since BET-bromodomain inhibitors have been shown to drastically reduce tumor growth in many cancer types including hematological malignancies and a variety of solid tumors[44].

FP-RMS and ES are two pediatric tumors driven by a fusion transcription factor, PAX3-FOXO1 and EWS-FLI1 respectively, generated by a chromosomal translocation. These aberrant TFs create and bind to enhancers and SEs, together with BRD4, from where they regulate gene expression related to cell proliferation, invasiveness, and cell survival[24,29,45,46]. Therefore, both tumors are transcriptionally addicted, and dependent on BRD4 to regulate their enhancer-driven expression signature which allows tumor growth[24–26,47,48]. Here, we demonstrate that FP-RMS and ES are both dependent on the chromatin remodeler CHD4 as well. We describe the detailed mechanisms by which CHD4 influences the viability of these pediatric sarcomas and define a new role for CHD4 in oncogenic transcription regulation. In FP-RMS, we demonstrate for the first time that CHD4 binds to enhancers and SEs together with BRD4 and the oncogenic driver P3F. The presence of CHD4 in SEs is essential to keep these regions open and permissive to the binding of P3F and BRD4, which is required for the expression of a subset of P3F-target genes (Fig.7). Similarly, in mouse embryonic stem cells, CHD4 was found to regulate the positioning of the TF Sall4[49]. Additionally, we found that CHD4 directly interacts with BRD4 which suggests that these epigenetic factors might collaborate for the regulation of SE-driven transcription. Moreover, our transcriptomic data after CHD4 depletion revealed that this chromatin remodeler surprisingly controls the expression of known EWS-FLI1 target genes in FP-RMS, which suggests a common gene signature among these tumors. These results can help to identify the fusion protein's target genes that are essential for tumor development.

CHD4 is commonly described as a core member of the NuRD complex, although recent reports describe it as a peripheral unit[37]. Our results show that CHD4 co-localizes with NuRD members, such as HDAC2 and RBBP4, at P3F-containing SEs, however CHD4 free RBBP4+HDAC2-containing regions were also observed. Moreover, BRD4 not only interacts with CHD4 but also with the NuRD members HDAC2 and MTA2. Hence, we investigated the role of the NuRD complex in FP-RMS cell viability and found that, besides CHD4 depletion, knockout of RBBP4, HDAC1, and GATAD2A/B also affected FP-RMS cell proliferation. However, the knockout of the MBD, which are core and mutually exclusive subunits of NuRD[50,51], and of the MTA proteins had no effect on FP-RMS viability, which supports the hypothesis that CHD4 acts independently of NuRD to maintain FP-RMS cell proliferation.

Furthermore, RBBP4 and HDAC1 are associated with other remodeling complexes like SIN3, coREST, PRC2 and CAF-1[39], hence, the effects observed upon their knockout might not reflect the actions of the NuRD complex. In addition, unlike CHD4 depletion, silencing of RBBP4 did not affect the expression of P3F target genes nor did it cause tumor cell death. Interestingly, in mouse embryonic stem cells, in the absence of nuclear Mbd3, Chd4 remains only associated with Gatad2b in the nucleus[52]. These findings suggest a supporting function of GATAD2B for CHD4 and could explain why we observe an effect upon GATAD2A/B depletion.

Besides FP-RMS and ES, published data has shown that CHD4 depletion affects the growth of a number of additional tumors like glioblastoma[53], liver cancer[54], lung cancer[55], and colorectal cancer[56,57]. Furthermore, and unlike other chromatin remodelers, CHD4 is rarely found mutated in cancer[58], which could be an indicator of tumor dependency. Our analysis of two publicly available databases of tumor susceptibilities, Achilles and DRIVE, identified CHD4, amongst the SNF2-like ATPases and the NuRD members, as a new tumor susceptibility. The data indicate that CHD4 depletion, similar to the one of BRD4, impaired tumor cell viability in the majority of the cancer cell lines analyzed in both projects. Intriguingly, tumor sensitivity to CHD4 depletion seems to be selective since silencing of this chromatin remodeler has no influence on the cell proliferation of myoblasts and fibroblasts[27]. Supporting this observation, a RNAi screen identified CHD4 as an essential gene in breast cancer cells with no effect on the cell proliferation of non-transformed mammary epithelial cells (MCF10A)[59]. Another study has also shown that CHD4 is required for the maintenance of childhood acute myeloid leukemia but not for the proliferation and survival of normal primary hematopoietic cells[60]. These studies demonstrate the potential of CHD4 as a cancer-specific therapeutic target. In this study, we describe a positive effect of CHD4 in SE-driven gene expression that might partially explain the broad tumor dependency to CHD4. Nevertheless, the mechanisms associated with CHD4 dependency may vary among tumors. Colorectal cancer depends on CHD4 since it mediates the silencing of tumor suppressor genes by recruitment of DNMTs and EZH2[56]. In acute myeloid leukemia, CHD4 depletion induces transcription alterations that impair colony formation[61]. CHD4 has also been associated with DNA-damage repair which might have implications in the genome stability of tumor cells. In fact, CHD4 depletion leads to the accumulation of spontaneous DNA damage, increased ionizing radiation sensitivity, and impaired homologous recombination repair[62,63].

In conclusion, by further investigating the role of CHD4 in FP-RMS, we describe for the first time the contribution of the remodeling activity of CHD4 for the control of SE-accessibility to oncogenic transcription factors and cofactors. Interestingly, CHD4 seems to control SE-

driven gene expression and tumor cell viability in a NuRD-independent manner. In ES, we observe that CHD4 is also indispensable for tumor cell viability, and, since it collaborates with EWS-FLI1 for its repressive activity[28], we hypothesize that this chromatin remodeler acts in a similar way in this tumor as in FP-RMS. Nevertheless, further studies are required to investigate the role of CHD4 in EWS-FLI1 activity. In addition, we identify a novel broad tumor dependency to CHD4 which can be partially explained by its influence in oncogenic SE-driven gene expression. Hence, this work puts CHD4 side-by-side with BRD4 as a prominent and promising new target for SE-disruption in cancer therapy.

MATERIAL AND METHODS

Cell lines

The FP-RMS cell line RH4 was kindly provided by Peter Houghton at the Research Institute of the Nationwide Children's Hospital, Columbus OH. The HEK293T cell line was used to produce lentiviral particles and was purchased from the American Type Cell Culture Collection, LGC Promochem, Molsheim Cedex, France. Both cell lines were routinely maintained in Dulbecco's modified Eagle's medium (DMEM, Sigma-Aldrich) supplemented with 10% fetal bovine serum (#F7524, Sigma Aldrich), 2 mM L-glutamine (Bioconcept AG) and 100 U/ml penicillin/streptomycin (Thermo Fisher Scientific AG). The Ewing sarcoma cell lines, A673 and SKNMC, were cultured in RPMI (Roswell Park Memorial Institute, Sigma-Aldrich) instead of DMEM. All cells were cultured in 5% CO_2 at 37°C. The RH4 cell line was tested and authenticated by cell line typing analysis (STR profiling) in 2014/2015 and positively matched[66].

CRISPR/Cas9 flag knockin

The knockin of a 3xflag peptide in endogenous CHD4 or BRD4 in RH4 cells was performed using the CRISPR/Cas9 technology[67]. Guide RNAs targeting the N- or C-terminal of CHD4 or the N-terminal of BRD4 were cloned into the pSpCas9(BB)-2A-Puro PX459 vector (#62988, addgene) using the appropriate primers and the In-Fusion cloning HD kit (Clontech Laboratories Inc., Mountain View, CA, USA) according to the manufacturer's instructions. Cloning was verified by sequencing. Then, the PX459 vector was transiently co-transfected with the specific donor DNA into RH4 cells using the JetPrime reagent (Polyplus Transfections, Illkirch, France). One day after transfection, cells were incubated with 1μM of non-homologous end-joining inhibitor SCR7. Clones were obtained by limited dilution and flag insertion was confirmed by immunofluorescence and western blot. Detailed information regarding the sequences of sgRNAs and donor DNAs can be found in Supplementary Tables 3 and 4.

Chromatin Immunoprecipitation-sequencing (ChIP-seq)

ChIP assays were performed by using the ChIP-IT High Sensitivity kit (#53040, Active Motif) according to the manufacturer's instructions. Briefly, cells were grown to confluence, fixed with 1% formaldehyde (Sigma-Aldrich) for 13 min, harvested and sonicated with the EpiShear™ ProbeSonicator (#53052, Active Motif) for 27 cycles (30% amp, 30sec ON, 30sec OFF). Sonicated lysates were then quantified and 30ug of chromatin were incubated overnight at 4°C with 5-10ug of antibody (antibodies listed in Supplementary Table 5). DNA was purified according to the manufacturer's instructions. Details regarding sequencing can be found in supplementary material and methods.

DNase Hypersensitivity assays

DNase hypersensitivity assays were performed as previously described[68]. Per condition, 10 000 cells were washed and resuspended in RSB buffer (10mM Tris-HCl pH 7.4, 10mM NaCl, 3mM $MgCl_2$, 0.2% Triton X-100). DNase I (#04-716-728-001, Roche; 10µL of a 0.33 U/µL DNase solution in RSB buffer) was added to a total of 40µL of cells and incubated for 5 minutes at 37 ˚C. The digestion was halted with 50 µL of stop buffer (10 µM Tris-HCl pH 7.4, 10 µM NaCl, 10µM EDTA, 0.15% SDS, supplemented with 125 µL proteinase K (#03115828001, Roche) per 10mL). Proteinase K activation at 55 ˚C for 1 hour was followed by DNA purification (NucleoSpin Gel and PCR Clean-up, Macherey-Nagel GmbH & Co. KG). Library preparation was performed as with ChIP-seq samples, except that paired-end was employed rather than single-end sequencing on the NextSeq 500 (Illumina).

Co-immunoprecipitation assays

Cells from three confluent 15cm dishes, per condition, were washed with PBS and harvested. Then, they were lysed with sucrose buffer (320mM sucrose, 3mM $CaCl_2$, 2mM MgOAc, 0.1 mM EDTA, 10mM DTT, 0.5mM PMSF, 0.25% NP-40) and nuclei were pelleted by centrifugation (10min, 1100g, 4°C). Cytoplasmic fraction was collected and stored at -20°C. After wash with NP-40 free sucrose buffer, nuclei were lysed by incubation with lysis buffer (50mM HEPES pH 7.8, 3mM $MgCl_2$, 300mM NaCl, 1mM DTT, 0.1mM PMSF) in presence of 15u/ul of benzonase for 1hr at 4°C. Before antibody incubation, an input sample was collected and stored at -20°C. Protein G Dynabeads were coupled with 8µg of anti-flag antibody (clone M2, #F1804, Sigma Aldrich) per plate used and then incubated overnight together with the nuclear extracts. After wash, immunoprecitates were eluted in 50mM Tris-HCl pH 7.4, 150mM NaCl and 200µg/ml of flag peptide (#F4799, Sigma Aldrich) and analyzed by western blotting.

NuRD-centered CRISPR screen

RH4 cells stably expressing Cas9 were obtained by transducing wildtype cells with the expression vector lentiCRISPRv2 puro (#98290, addgene) followed by puromycin selection (1µg/mL). Cas9 expression was confirmed by western blot and immunofluorescence (data not shown).

Guide RNAs targeting the NuRD members or BRD4 were cloned into the RFP-labelled sgRNA expression construct (named here shuttle vector), and the control guide targeting the AAVS1 locus into the BFP-labelled shuttle vector. Both shuttle vectors were kindly provided by Dr. Yun Huang, University Children's Hospital Zurich. Viruses were produced in HEK293T cells by co-transfection of pVSV-G, PAX2, and the shuttle vectors using $CaPO_4$. Medium was replaced 24hrs after transfection and viruses were harvested after additional 48hrs. Viral supernatant was cleared by centrifugation, filtered and concentrated (Amicon® Ultra 15 mL, Millipore).

Cas9 expressing RH4 cells were infected with the viral supernatant, supplemented with 8µg/ml polybrene (Sigma Aldrich) for 24hrs, containing either RFP-labelled or BFP-labelled lentiviral shuttle vectors. Two days after transduction the red and blue populations were mixed 1:1. The distribution of red and blue cells was assessed by FACS analysis, 2 and 12 days after transduction. The results are shown as the ratio between red and blue population at day 12 normalized to day 2. Five guides per target were used. The sequences regarding the guide RNAs used can be found in Supplementary Table 6.

Doxycycline-inducible knockdowns

The pRSIT-U6Tet-shRNA-PGKTetRep-2A-GFP-2A-puro vectors with 2 shRNAs against RBBP4 were purchased from Cellecta Inc. with the following target sequences: shRBBP4#1 5' GCCTTTCTTTCAATCCTTATA 3' and shRBBP4#2 5' e: ATGAACCTTGGGTGATTTGTT 3'. Viruses were produced by co-transfection of the shRNA vectors and lentiviral packaging and envelope plasmids (pMDL, pREV and pVSV-G kindly provided by Oliver Pertz, Department of Biomedicine, University of Basel, Switzerland) into HEK293Tcells using $CaPO_4$. Medium was replaced 24hrs after transfection and viruses were harvested after additional 48hrs. Viral supernatant was cleared by centrifugation, filtered and concentrated (Amicon® Ultra 15 mL, Millipore). RH4 cells were transduced for 24 hours in the presence of 8µg/ml polybrene (Sigma-Aldrich) and selected with puromycin (1µg/mL). RH4 scramble control RH4 cells carrying CHD4 or PAX3-FOXO1 targeting shRNAs were previously established[27]. ShRNA expression was induced using 100ng/ml doxycycline (Sigma-Aldrich). Data were normalized to corresponding uninduced cells and compared to scramble shRNA expressing cells.

Xenograft studies

Five million SKNMC cells, containing stably integrated doxycycline-inducible shCHD4#1 or shScramble expression constructs, were engrafted subcutaneously in 30 mice (15 mice per construct). NOD/Scid female mice of 8 to 12 weeks old were used for this experiment (Charles River, Sulzfeld, Germany). Mice were divided in treatment (8 mice) and vehicle (7 mice) groups, and, once the tumors reached an average of 60mm^3/group, they were treated intraperitoneally with either sterile PBS or a doxycycline solution (53.3 mg/kg) for two consecutive days. Additionally, starting from the first day of treatment, mice were fed with doxycycline-supplemented food (625 mg doxycycline/kg) or control food. Tumor size was determined by daily measurements of two diameters (d1, d2) in right angles using a digital caliper. Total tumor volumes were calculated using the following formula ($V= (4/3) \pi r3$; $r= (d1+d2)/4$). Mice were euthanized when the tumors reached a volume of 1000 mm^3. All animal experiments have been approved by the Swiss veterinary authorities and were performed according to the animal license ZH206/15.

Statistics

All statistical analysis (with the exception of GSEA analysis) were performed with the GraphPad prism software, version 7. Data are represented as mean ± SD, unless otherwise noted.

RNA-seq

For gene expression profiling, RNA of RH4 cells, containing either a doxycycline-inducible scramble shRNA or shRNAs targeting CHD4 and PAX3-FOXO1, was isolated 24 and 48hrs after the start of doxycycline treatment using RNeasy Mini Kit (Qiagen), including the DNase digestion step, according to the manufacturer`s instructions. For more details regarding sequencing and data analysis please refer to the supplementary material and methods.

ACKNOWLEDGEMENTS

We are very grateful to Dr. Marielle Yohe of the NCI-NIH for her support and helpful discussions. Also, we would like to thank Dr. Yun Huang of the University Children's Hospital Zurich for providing us with the shuttle sgRNA plasmid.

REFERENCES

1. Weinstein, I. B. & Joe, A. K. Mechanisms of Disease: oncogene addiction—a rationale for molecular targeting in cancer therapy. *Nat. Clin. Pract. Oncol.* **3,** 448–457 (2006).
2. Bradner, J. E., Hnisz, D. & Young, R. A. Transcriptional Addiction in Cancer. *Cell* **168,**

629–643 (2017).

3. Koumenis, C. & Giaccia, A. Transformed cells require continuous activity of RNA polymerase II to resist oncogene-induced apoptosis. *Mol. Cell. Biol.* **17,** 7306–16 (1997).

4. Chen, Z., Li, S., Subramaniam, S., Shyy, J. Y.-J. & Chien, S. Epigenetic Regulation: A New Frontier for Biomedical Engineers. *Annu. Rev. Biomed. Eng.* **19,** 195–219 (2017).

5. Plass, C. *et al.* Mutations in regulators of the epigenome and their connections to global chromatin patterns in cancer. *Nat. Rev. Genet.* **14,** 765–780 (2013).

6. Yoo, C. B. & Jones, P. A. Epigenetic therapy of cancer: past, present and future. *Nat. Rev. Drug Discov.* **5,** 37–50 (2006).

7. Dhanak, D. & Jackson, P. Development and classes of epigenetic drugs for cancer. *Biochem. Biophys. Res. Commun.* **455,** 58–69 (2014).

8. Helin, K. & Dhanak, D. Chromatin proteins and modifications as drug targets. *Nature* **502,** 480–488 (2013).

9. Jang, M. K. *et al.* The Bromodomain Protein Brd4 Is a Positive Regulatory Component of P-TEFb and Stimulates RNA Polymerase II-Dependent Transcription. *Mol. Cell* **19,** 523–534 (2005).

10. Devaiah, B. N. *et al.* BRD4 is an atypical kinase that phosphorylates Serine2 of the RNA Polymerase II carboxy-terminal domain. *Proc. Natl. Acad. Sci.* **109,** 6927–6932 (2012).

11. Yokoyama, Y. *et al.* BET Inhibitors Suppress ALDH Activity by Targeting *ALDH1A1* Super-Enhancer in Ovarian Cancer. *Cancer Res.* **76,** 6320–6330 (2016).

12. Bhagwat, A. S. *et al.* BET Bromodomain Inhibition Releases the Mediator Complex from Select cis -Regulatory Elements. *Cell Rep.* **15,** 519–530 (2016).

13. Chapuy, B. *et al.* Discovery and characterization of super-enhancer-associated dependencies in diffuse large B cell lymphoma. *Cancer Cell* **24,** 777–90 (2013).

14. Delmore, J. E. *et al.* BET bromodomain inhibition as a therapeutic strategy to target c-Myc. *Cell* **146,** 904–17 (2011).

15. Clapier, C. R. & Cairns, B. R. The Biology of Chromatin Remodeling Complexes. *Annu. Rev. Biochem.* **78,** 273–304 (2009).

16. Clapier, C. R., Iwasa, J., Cairns, B. R. & Peterson, C. L. Mechanisms of action and regulation of ATP-dependent chromatin-remodelling complexes. *Nat. Rev. Mol. Cell Biol.* **18,** 407–422 (2017).

17. Flaus, A., Martin, D. M. A., Barton, G. J. & Owen-Hughes, T. Identification of multiple distinct SNF2 subfamilies with conserved structural motifs. *Nucleic Acids Res.* **34,** 2887–2905 (2006).

18. Shern, J. F. *et al.* Comprehensive Genomic Analysis of Rhabdomyosarcoma Reveals a Landscape of Alterations Affecting a Common Genetic Axis in Fusion-Positive and Fusion-Negative Tumors. *Cancer Discov.* **4,** 216–231 (2014).

19. De Giovanni, C., Landuzzi, L., Nicoletti, G., Lollini, P.-L. & Nanni, P. Molecular and cellular biology of rhabdomyosarcoma. *Futur. Oncol.* **5,** 1449–1475 (2009).

20. Brohl, A. S. *et al.* The Genomic Landscape of the Ewing Sarcoma Family of Tumors

Reveals Recurrent STAG2 Mutation. *PLoS Genet.* **10,** e1004475 (2014).

21. Tirode, F. *et al.* Genomic landscape of Ewing sarcoma defines an aggressive subtype with co-association of STAG2 and TP53 mutations. *Cancer Discov.* **4,** 1342–53 (2014).

22. Kovar, H. *et al.* EWS/FLI-1 antagonists induce growth inhibition of Ewing tumor cells in vitro. *Cell Growth Differ.* **7,** 429–37 (1996).

23. Bernasconi, M., Remppis, A., Fredericks, W. J., Rauscher, F. J. & Schäfer, B. W. Induction of apoptosis in rhabdomyosarcoma cells through down-regulation of PAX proteins. *Proc. Natl. Acad. Sci. U. S. A.* **93,** 13164–9 (1996).

24. Gryder, B. E. *et al.* PAX3-FOXO1 Establishes Myogenic Super Enhancers and Confers BET Bromodomain Vulnerability. *Cancer Discov.* **7,** 884–899 (2017).

25. Hensel, T. *et al.* Targeting the EWS-ETS transcriptional program by BET bromodomain inhibition in Ewing sarcoma. *Oncotarget* **7,** 1451–63 (2016).

26. Jacques, C. *et al.* Targeting the epigenetic readers in Ewing sarcoma inhibits the oncogenic transcription factor EWS/Fli1. *Oncotarget* **7,** 24125–40 (2016).

27. Böhm, M. *et al.* Helicase CHD4 is an epigenetic coregulator of PAX3-FOXO1 in alveolar rhabdomyosarcoma. *J. Clin. Invest.* **126,** 4237–4249 (2016).

28. Sankar, S. *et al.* Mechanism and relevance of EWS/FLI-mediated transcriptional repression in Ewing sarcoma. *Oncogene* **32,** 5089–100 (2013).

29. Cao, L. *et al.* Genome-wide identification of PAX3-FKHR binding sites in rhabdomyosarcoma reveals candidate target genes important for development and cancer. *Cancer Res.* **70,** 6497–508 (2010).

30. Torrado, M. *et al.* Refinement of the subunit interaction network within the nucleosome remodelling and deacetylase (NuRD) complex. *FEBS J.* **284,** 4216–4232 (2017).

31. Ernst, J. *et al.* Mapping and analysis of chromatin state dynamics in nine human cell types. *Nature* **473,** 43–49 (2011).

32. Blake, J. A. & Ziman, M. R. Pax genes: regulators of lineage specification and progenitor cell maintenance. *Development* **141,** 737–51 (2014).

33. Hnisz, D. *et al.* Super-enhancers in the control of cell identity and disease. *Cell* **155,** 934–47 (2013).

34. Sengupta, S. & George, R. E. Super-Enhancer-Driven Transcriptional Dependencies in Cancer. *Trends in Cancer* **3,** 269–281 (2017).

35. Wai, D. C. C. *et al.* The BRD3 ET domain recognizes a short peptide motif through a mechanism that is conserved across chromatin remodelers and transcriptional regulators. *J. Biol. Chem.* **293,** 7160–7175 (2018).

36. Xue, Y. *et al.* NURD, a Novel Complex with Both ATP-Dependent Chromatin-Remodeling and Histone Deacetylase Activities. *Mol. Cell* **2,** 851–861 (1998).

37. Low, J. K. K. *et al.* CHD4 Is a Peripheral Component of the Nucleosome Remodeling and Deacetylase Complex. *J. Biol. Chem.* **291,** 15853–66 (2016).

38. Hoffmeister, H. *et al.* CHD3 and CHD4 form distinct NuRD complexes with different yet overlapping functionality. *Nucleic Acids Res.* **45,** 10534–10554 (2017).

39. Basta, J. & Rauchman, M. The Nucleosome Remodeling and Deacetylase (NuRD)

Complex in Development and Disease. doi:10.1016/j.trsl.2014.05.003

40. Taylor, M. V. & Hughes, S. M. Mef2 and the skeletal muscle differentiation program. *Semin. Cell Dev. Biol.* **72,** 33–44 (2017).

41. Nitarska, J. *et al.* A Functional Switch of NuRD Chromatin Remodeling Complex Subunits Regulates Mouse Cortical Development. *Cell Rep.* **17,** 1683–1698 (2016).

42. Wilczewski, C. M. *et al.* CHD4 and the NuRD complex directly control cardiac sarcomere formation. *Proc. Natl. Acad. Sci.* 201722219 (2018). doi:10.1073/pnas.1722219115

43. McDonald, E. R. *et al.* Project DRIVE: A Compendium of Cancer Dependencies and Synthetic Lethal Relationships Uncovered by Large-Scale, Deep RNAi Screening. *Cell* **170,** 577–592.e10 (2017).

44. Erez-Salvia, M. P. & Esteller, M. Bromodomain inhibitors and cancer therapy: From structures to applications. *Epigenetics* **12,** (2017).

45. Riggi, N. *et al.* EWS-FLI1 utilizes divergent chromatin remodeling mechanisms to directly activate or repress enhancer elements in Ewing sarcoma. *Cancer Cell* **26,** 668–681 (2014).

46. Boulay, G. *et al.* Cancer-Specific Retargeting of BAF Complexes by a Prion-like Domain. *Cell* **171,** 163–178.e19 (2017).

47. Loganathan, S. N. *et al.* BET bromodomain inhibitors suppress EWS-FLI1-dependent transcription and the IGF1 autocrine mechanism in Ewing sarcoma. *Oncotarget* **7,** 43504–43517 (2016).

48. Gollavilli, P. N. *et al.* EWS/ETS-driven Ewing Sarcoma requires BET bromodomain proteins. *Cancer Res.* canres.0484.2018 (2018). doi:10.1158/0008-5472.CAN-18-0484

49. Miller, A. *et al.* Sall4 controls differentiation of pluripotent cells independently of the Nucleosome Remodelling and Deacetylation (NuRD) complex. *Development* **143,** 3074–84 (2016).

50. Le Guezennec, X. *et al.* MBD2/NuRD and MBD3/NuRD, Two Distinct Complexes with Different Biochemical and Functional Properties. *Mol. Cell. Biol.* **26,** 843–851 (2006).

51. Zhang, W. *et al.* The Nucleosome Remodeling and Deacetylase Complex NuRD Is Built from Preformed Catalytically Active Sub-modules. *J. Mol. Biol.* **428,** 2931–2942 (2016).

52. Bornelöv, S. *et al.* The Nucleosome Remodeling and Deacetylation Complex Modulates Chromatin Structure at Sites of Active Transcription to Fine-Tune Gene Expression. *Mol. Cell* **71,** 56–72.e4 (2018).

53. Chudnovsky, Y. *et al.* ZFHX4 Interacts with the NuRD Core Member CHD4 and Regulates the Glioblastoma Tumor-Initiating Cell State. *Cell Rep.* **6,** 313–324 (2014).

54. Nio, K. *et al.* Defeating EpCAM+ liver cancer stem cells by targeting chromatin remodeling enzyme CHD4 in human hepatocellular carcinoma. *J. Hepatol.* **63,** 1164–1172 (2015).

55. Xu, N., Liu, F., Zhou, J. & Bai, C. CHD4 is associated with poor prognosis of non-small cell lung cancer patients through promoting tumor cell proliferation. in *11.1 Lung Cancer* **48,** PA2862 (European Respiratory Society, 2016).

56. Xia, L. *et al.* CHD4 Has Oncogenic Functions in Initiating and Maintaining Epigenetic Suppression of Multiple Tumor Suppressor Genes. *Cancer Cell* **31,** 653–668.e7 (2017).

57. Cai, Y. *et al.* The NuRD complex cooperates with DNMTs to maintain silencing of key colorectal tumor suppressor genes. *Oncogene* **33,** 2157–2168 (2014).

58. Le Gallo, M. *et al.* Exome sequencing of serous endometrial tumors identifies recurrent somatic mutations in chromatin-remodeling and ubiquitin ligase complex genes. *Nat. Genet.* **44,** 1310–1315 (2012).

59. D'Alesio, C. *et al.* RNAi screens identify CHD4 as an essential gene in breast cancer growth. *Oncotarget* **7,** 80901–80915 (2016).

60. Heshmati, Y. *et al.* Identification of CHD4 As a Potential Therapeutic Target of Acute Myeloid Leukemia. *Blood* **128,** (2016).

61. Sperlazza, J. *et al.* Depletion of the chromatin remodeler CHD4 sensitizes AML blasts to genotoxic agents and reduces tumor formation. *Blood* **126,** 1462–1472 (2015).

62. Qi, W. *et al.* Acetyltransferase p300 collaborates with chromodomain helicase DNA-binding protein 4 (CHD4) to facilitate DNA double-strand break repair. *Mutagenesis* **31,** 193–203 (2016).

63. Smeenk, G. *et al.* The NuRD chromatin–remodeling complex regulates signaling and repair of DNA damage. *J. Cell Biol.* **190,** 741–749 (2010).

FIGURES

Figure 1

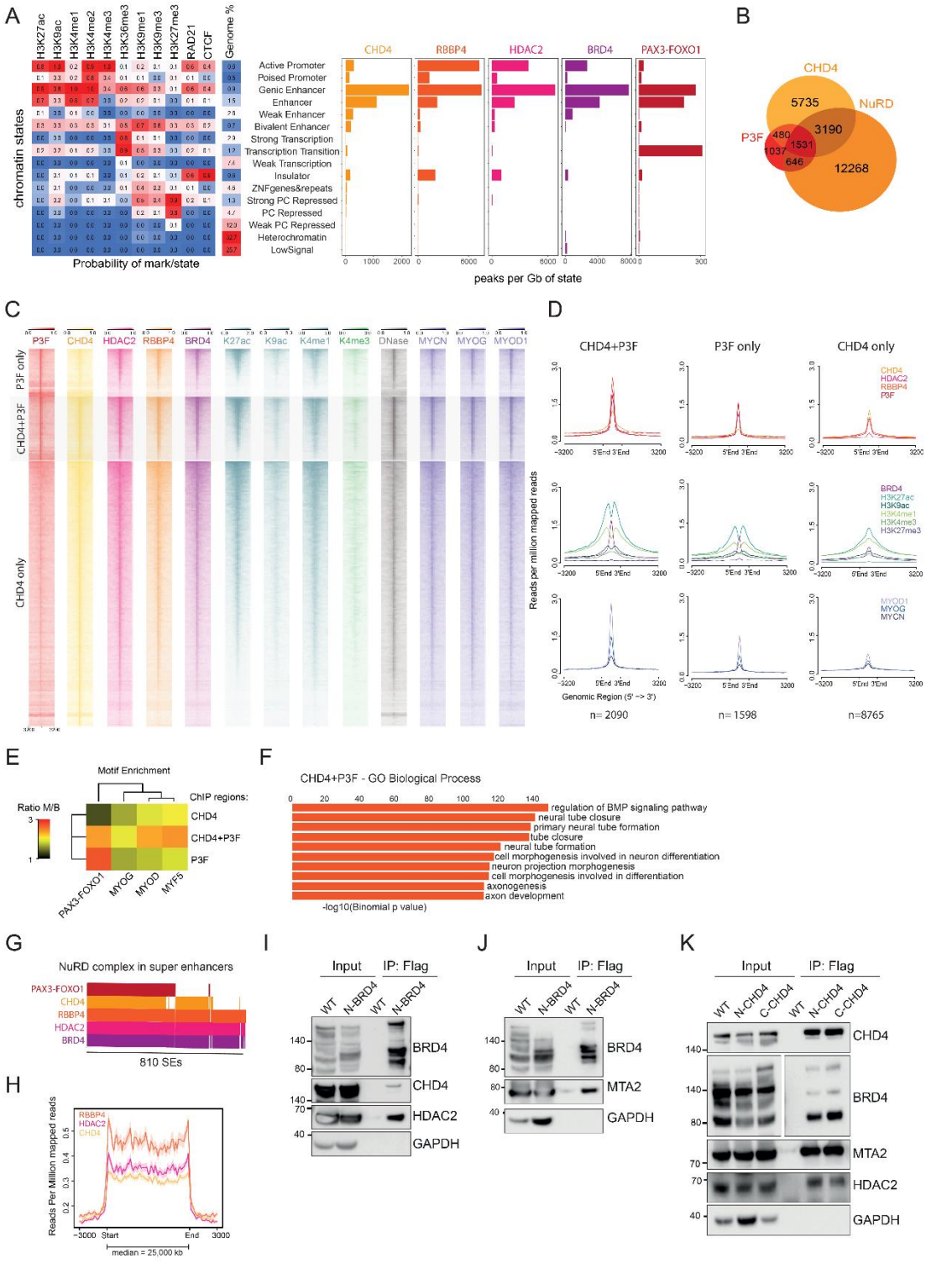

Fig.1 – CHD4 co-localizes with PAX3-FOXO1 at super-enhancers and interacts with BRD4
(A) Chromatin states identified in the FP-RMS cell line RH4 and abundance of ChIP-seq peaks/Gb of state of the indicated proteins. (B) Overlap of the ChIP-seq signal for PAX3-FOXO1 (P3F), CHD4 and NuRD (represented by the intersection between HDAC2 and RBBP4 peaks). (C) Heatmaps depicting the ChIP signal of the indicated proteins in RH4 cells at P3F-only (n=1598), CHD4+P3F (n=2090) and CHD4-only regions (n=8765). The rows show 7kb regions, centered on P3F or CHD4 peaks and ranked by the ChIP signal intensity of H3K27ac. Color shading corresponds to ChIP-seq read counts.

(D) Density plots displaying the average signal of the indicated proteins in the genome regions described in C. (E) Motif enrichment analysis performed with P3F-only, CHD4+P3F and CHD4-only locations. Heatmap depicts the ratio between M (the percentage of a given motif in the studied regions) and B (the percentage of the same motif in the background genome). (F) Gene ontology analysis (assigned by GREAT) of the biological processes associated with the CHD4+P3F regions. (G) Representative plot of the occupancy of each of the indicated proteins in the 810 super-enhancers (SEs) identified in RH4 cells. (H) Density plot showing the average signal of the NuRD members RBBP4, HDAC2, and CHD4 at SEs. (I-K) CoIP assays performed by using an anti-flag antibody to immunoprecipitate the endogenous N-terminal flag tagged BRD4 and the N- or C-terminal flag tagged CHD4. RH4 wildtype cells (WT) were used as negative control.

Figure 2

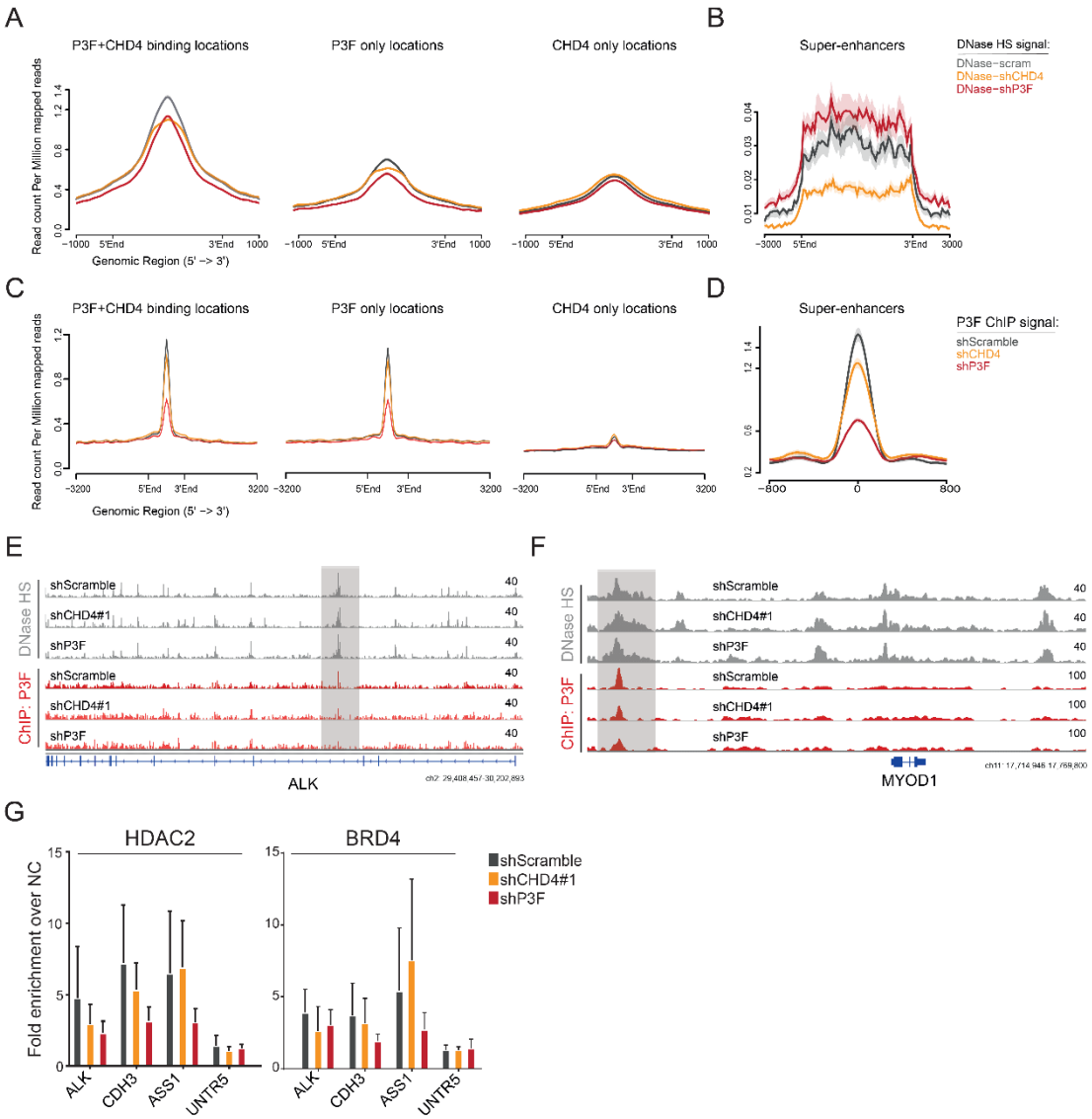

Fig.2 – CHD4 changes chromatin accessibility and allows P3F and BRD4 binding to SEs
(A/B) Density plots depicting the average DNase hypersensitivity (DNase HS) signal in RH4 cells upon CHD4 or P3F knockdown (48 hours of incubation with doxycycline). Each plot displays the DNase HS signal at CHD4+P3F (n=2090), P3F-only (n=1598), CHD4-only regions (n=8765), and at SEs (n=810). The signal shown is centered on P3F or CHD4 peaks. (C/D) Density plots depicting the average P3F ChIP-seq signal in RH4 cells upon CHD4 or P3F knockdown (48 hours of incubation with doxycycline). Each plot displays the P3F ChIP-seq signal at CHD4+P3F, P3F-only, CHD4-only regions, and at SEs. The signal shown is centered on P3F or CHD4 peaks. (E/F) Example of gene tracks of the DNase HS signal and of the P3F ChIP-seq signal at the ALK or MYOD1 SE loci in RH4 cells. (G) ChIP-

qPCR experiments show the decrease in HDAC2 and BRD4 binding in selected P3F-binding regions upon CHD4 or P3F silencing (48 hours of incubation with doxycycline). Data is displayed as fold enrichment over negative control region (n=2). UNTR5 served as negative control.

Figure 3

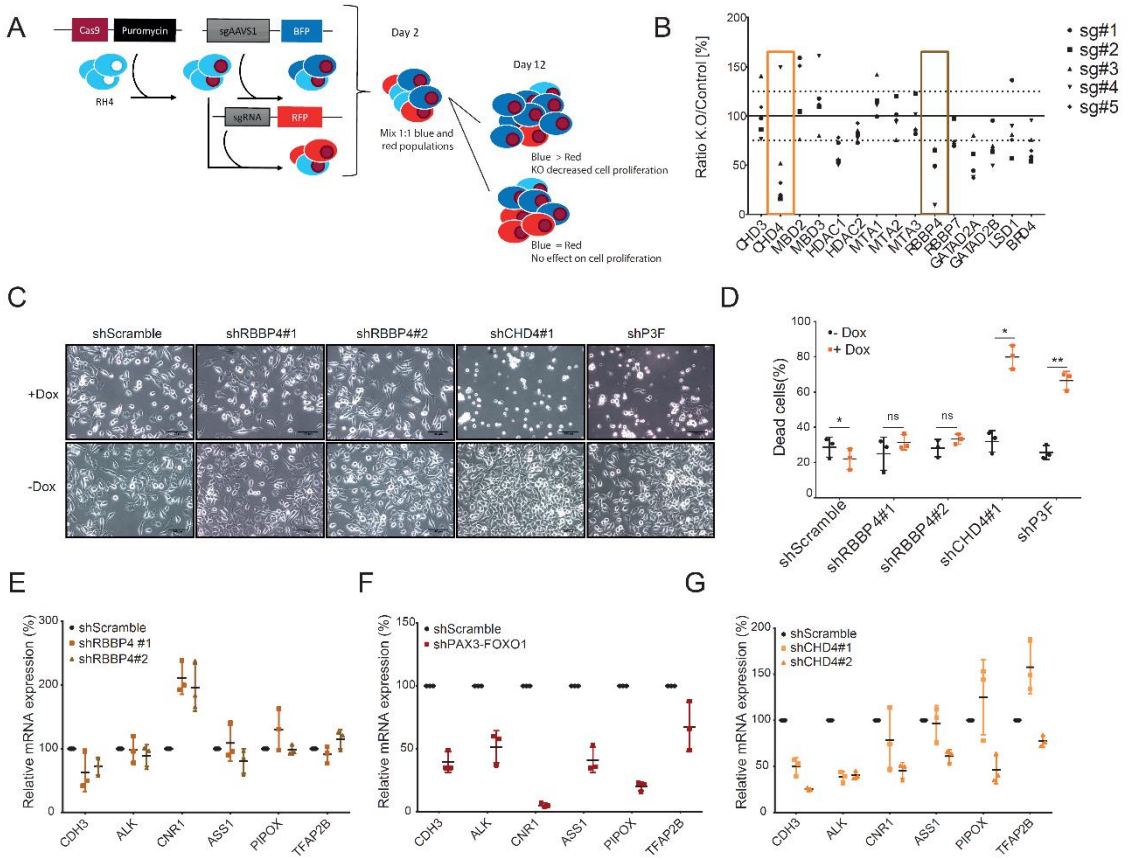

Fig.3 – NuRD-centered CRISPR screen reveals FP-RMS cell proliferation susceptibilities
(A) Illustrative scheme of the NuRD-centered CRISPR screen. (B) CRISPR screen results are displayed as the ratio, in percentage, between the NuRD knockout population and the control population at day 12, normalized to day 2. Each point represents the average of three biological replicates for each sgRNA used. Five sgRNAs were used per NuRD member. Guide RNAs targeting BRD4 and CHD4 were used as positive controls (C) Representative phase-contrast images of RH4 cells 5 days after doxycycline induction of shRNAs targeting RBBP4, CHD4, and P3F. A scramble shRNA was used as negative control. Also depicted are the uninduced RH4 cells carrying each construct. Scale bar - 100µm. (D) Percentage of dead cells observed 5 days after doxycycline-induction of shRNAs targeting RBBP4, CHD4 and P3F in RH4 cells. Cells were stained with 7AAD and analyzed by FACS. A scrambled shRNA was used as negative control. Values represented as mean ± SD of 3 independent experiments (*p< 0.1, **p < 0.01; ***p < 0.001, ratio paired t test) (E-G) Expression levels of the indicated P3F-target genes quantified by qPCR in RH4 cells 48hrs upon RBBP4, P3F and CHD4 knockdown induction with doxycycline. Expression levels (relative to GAPDH) were normalized to uninduced cells. Data represented as mean ± SD (n=3).

Figure 4

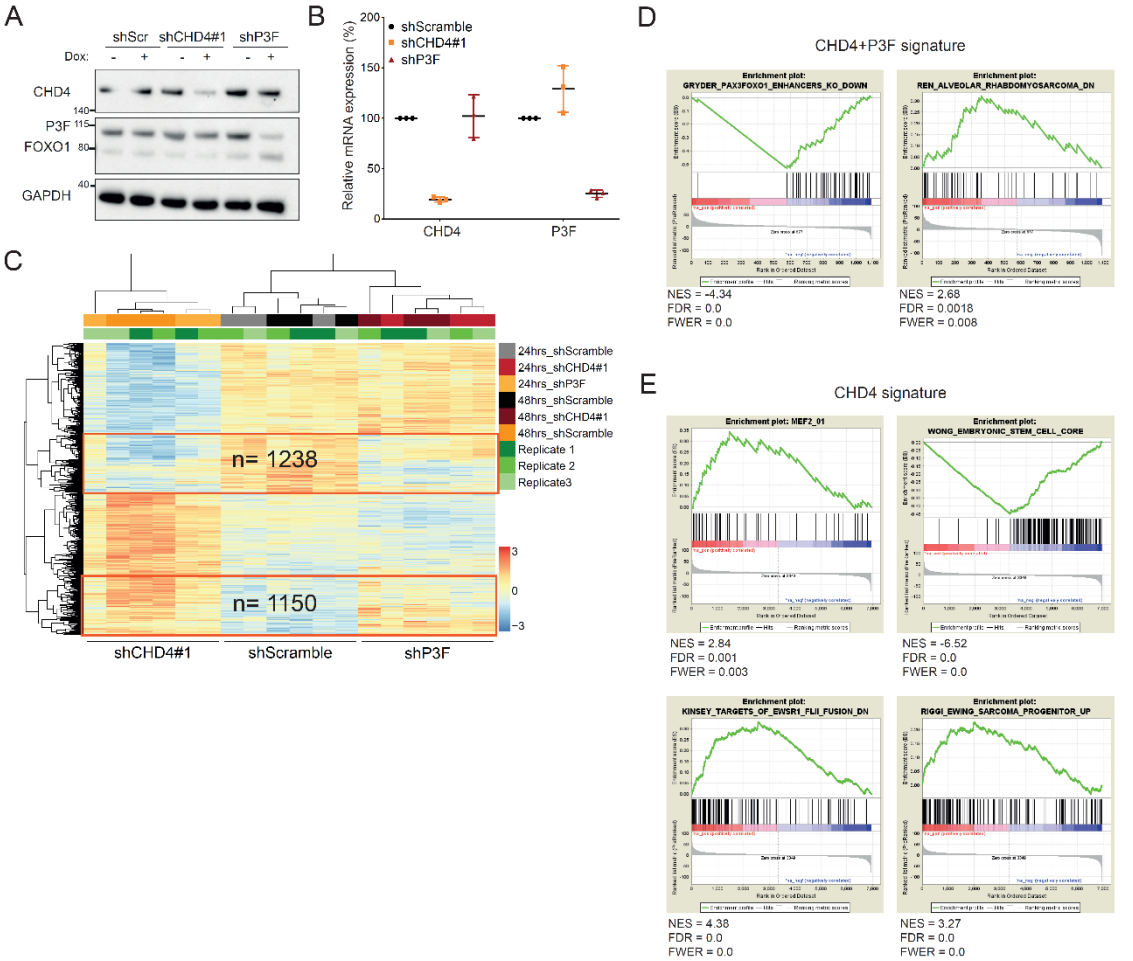

Fig.4 – CHD4 gene expression signature reveals common PAX3-FOXO1 and EWS-FLI1 target genes

(A) Western blot validates the knockdown obtained after 48hrs of doxycycline-induction of shRNAs targeting CHD4 or P3F. GAPDH was used as a loading control. (B) Expression levels of CHD4 and P3F quantified by qPCR in RH4 cells after 48hrs of silencing induction. Expression levels (relative to GAPDH) were normalized to uninduced cells. Data represented as mean ± SD (n=3). (C) Heatmap of unsupervised hierarchical clustering analysis of CHD4 and P3F knockdown profiles in RH4 cells after 24 or 48hrs of doxycycline-induced silencing. Displayed are the genes differentially expressed upon CHD4 silencing. A scrambled shRNA was used as control and 3 independent biological replicates were performed for each time point. (D/E) GSEA performed using either CHD4+P3F coregulated signature or CHD4 signature as preranked datasets. NES – normalized enrichment score, FDR – false discovery rate, FWER – family-wise error rate.

Figure 5

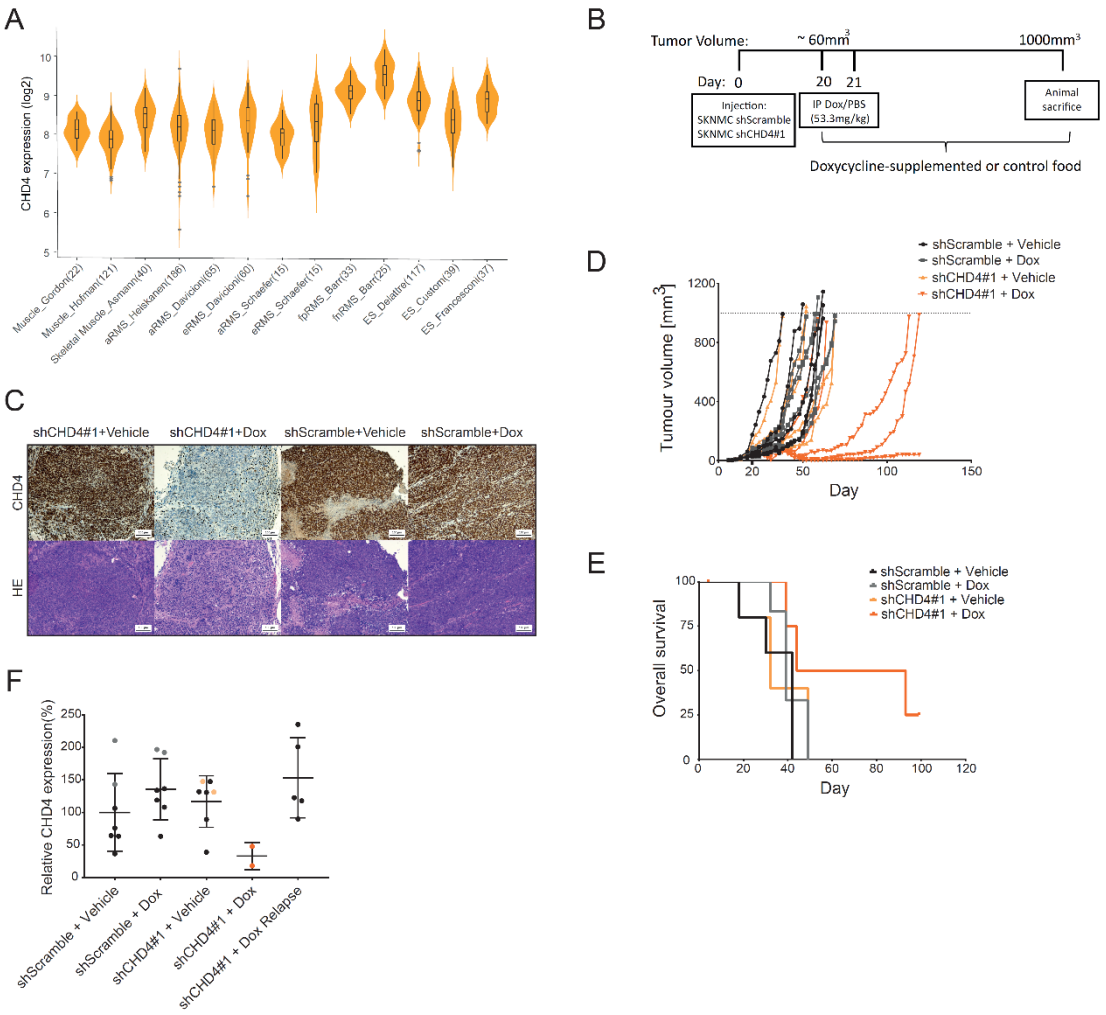

Fig.5 – CHD4 silencing impairs Ewing sarcoma tumor growth in vivo
(A) Violin plot displays CHD4 gene expression in the indicated datasets (R2 database). (B) Representative scheme of animal experiment. NOD/SCID mice were subcutaneously engrafted with SKNMC cells stably expressing a doxycycline-inducible shRNA targeting CHD4 or a scramble control. Mice bearing tumors of an average of 60mm^3 per group were treated intraperitoneally (IP) with a doxycycline or vehicle solution (PBS) for two consecutive days. From then on, mice were fed with doxycycline-supplemented food or control food. When tumors reached 1000mm^3, the animals were sacrificed. (C) Immunohistochemical staining of tumors with CHD4 antibody and HE (hematoxylin and eosin). Tumors were isolated 4 days after the beginning of the doxycycline treatment. Scale bar - 100µm. (D) Absolute tumor volumes measured by caliper at the indicated time points. Day 20 represents the start of the doxycycline treatment. Vehicle groups were composed of 5 mice and treated groups of 6. (E) Kaplan-Meier plot displays the overall survival of the different groups. (F) CHD4 mRNA expression level, quantified by qPCR, at the indicated groups. Each point represents one tumor isolated from one mouse. In black are displayed tumors isolated at the end of the experiment and in grey or orange are tumors isolated 4 days after the beginning of the doxycycline treatment.

Figure 6

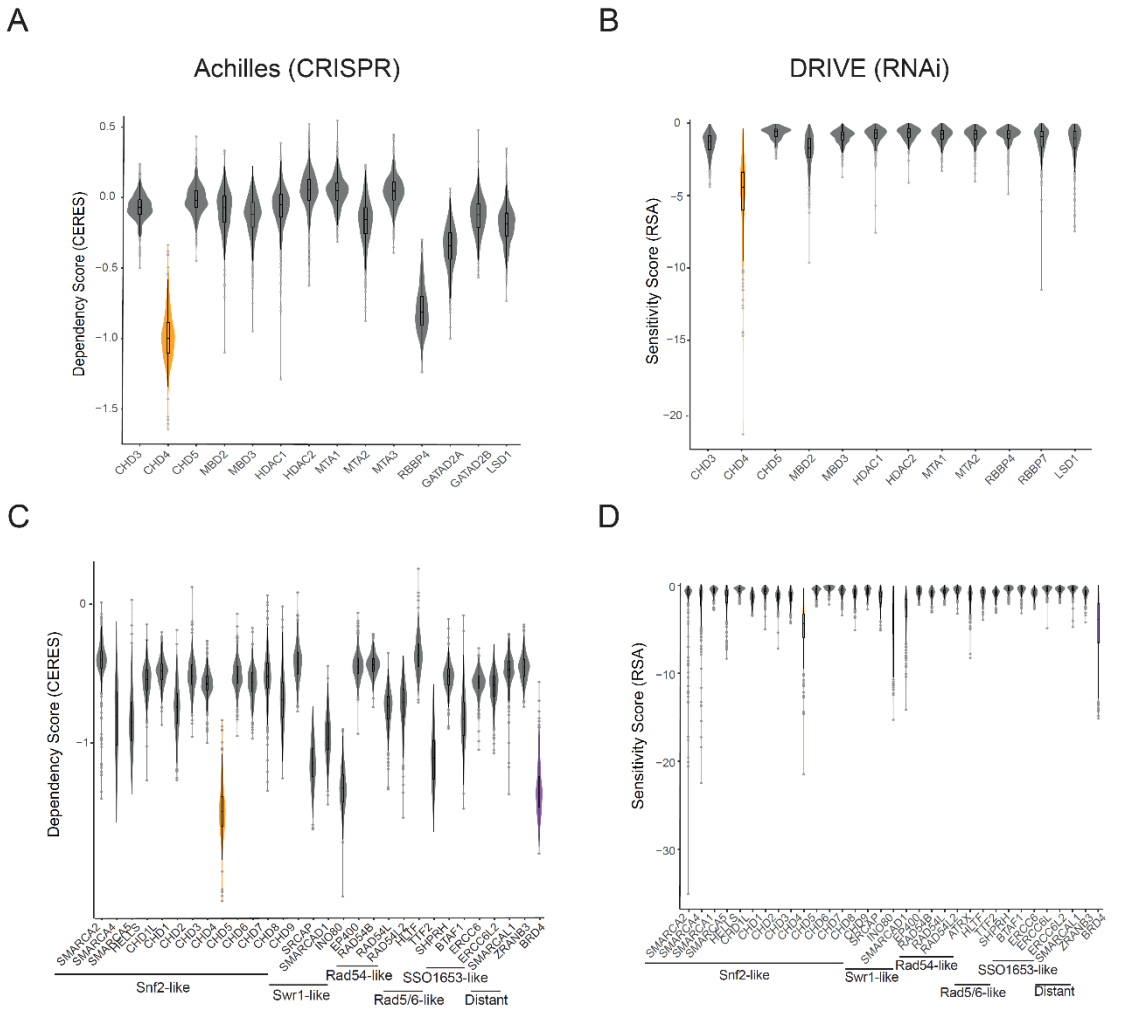

Fig.6 – CHD4 as a broad tumor dependency.
(A/C) Violin plots show the dependency scores of 391 cancer cell lines, calculated by CERES, for the indicated NuRD members and the chromatin remodelers of the SNF2-like family (Avana_public_18Q1 dataset). CHD4 is indicated in orange and BRD4 in purple. (B/D) Violin plots show the sensitivity score of 398 cancer cell lines, as RSA, for the indicated NuRD members and the chromatin remodelers of SNF2-like family (DRIVE platform). CHD4 is indicated in orange and BRD4 in purple.

Figure 7

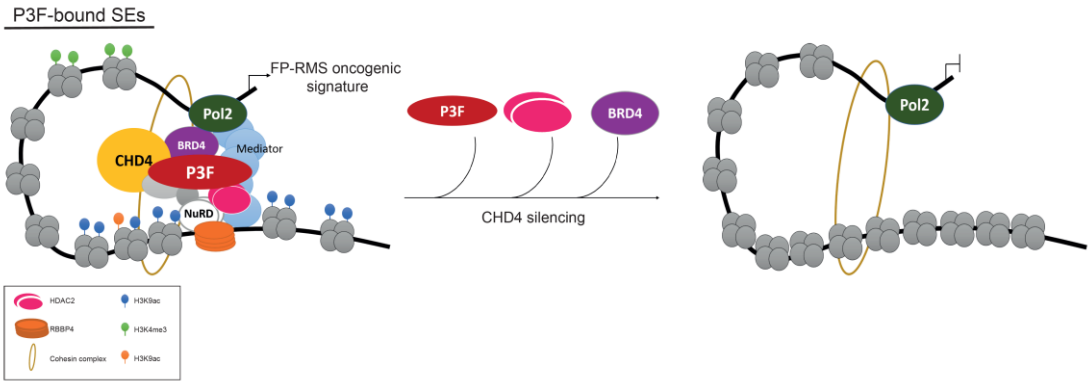

Fig.7 – CHD4 regulates SE accessibility to oncogenic TFs and cofactors.
In the presence of CHD4, P3F binds to SE, recruits many transcription cofactors, such as BRD4, and activates gene expression. By genetic depletion of CHD4, the P3F-bound SEs become less accessible and P3F and the cofactors BRD4 and HDAC2 are displaced from the chromatin which impairs SE-driven gene expression.

SUPPLEMENTARY MATERIAL AND METHODS

Supplementary Table 1 Tumor cell lines per tumor type tested in the Achilles database.

Tumor type	Number of cell lines
AML	10
Bladder	20
Breast	26
Cervical	1
Colon	25
Endometrial	14
Esophageal	10
Ewing's	4
Gastric	11
Glioblastoma	20
Glioma	10
Head and Neck	12
Kidney	17
Liver	13
Lung	48
Lymphoma	8
Medulloblastoma	7
Mesothelioma	1
Myeloma	7
Neuroblastoma	15
Osteosarcoma	5
Other	10
Ovarian	31
Pancreas	22
Rhabdoid	3
RMS	6
Sarcoma	1
Skin	29
Synovial Sarcoma	2
T-ALL	2
Thyroid	1
Total	**391**

Supplementary Table 2 *T*umor cell lines per tumor type, according with the pathologist annotation, tested in the DRIVE database.

Tumor Type	Number of cell lines
Biliary_Tract:Carcinoma	1
Bone:Sarcoma_Chondro	1
Cervix:Carcinoma	1
CNS:Medulloblastoma	1
Endometrium:Others	1
Liver:Carcinoma_Adeno	1
Liver:Hepatoblastoma	1
Lung:Others	1
PNET:Rhabdoid	1
Salivary_Gland:Carcinoma	1
Soft_Tissue:Sarcoma_Fibro	1
Undefined:GIST	1
Bone:Sarcoma_Osteo	2
Prostate:Carcinoma	2
Eye:Melanoma	3
Bone:Sarcoma_Ewing	4
Lymphoma:Multiple_Myeloma	4
Lymphoma:NH_B_cell	4
Leukemia:ALL	5
Lung:Mesothelioma	5
Soft_Tissue:Sarcoma_Synovial	5
Thyroid:Carcinoma	5
CNS:Glioma	6
Lung:NSCLC_Large_Cell	6
Lung:NSCLC_Others	6
Oesophagus:Carcinoma	8
Bladder:Carcinoma	9
PNET:Neuroblastoma	9
Soft_Tissue:Sarcoma_Rhabdoid	9
Upper_Aerodigestive_Tract:Carcinoma	9
Leukemia:AML	10
Lung:NSCLC_Squamous	10
Lung:SCLC	10
Endometrium:Carcinoma	11
CNS:Glioma_HighGrade	12
Liver:HCC	13
Kidney:Carcinoma	14
Gastric:Carcinoma	15
Ovary:Carcinoma	15
Pancreas:Carcinoma	22
Breast:Carcinoma	24
Skin:Melanoma	29
Lung:NSCLC_Adeno	30
Colorectal:Carcinoma	40
Total	**368**

Supplementary Table 3 Guide RNA sequences used in flag knockins.

Primer Name	Sequence
sg_NCHD4	GAGCGGAAGGGGATGGCGTC
sg_CCHD4	TCTGCATCTTCACTGCTGCT
sg_NBRD4	ATGTCTGCGGAGAGCGGCCCTGG

Supplementary Table 4 Donor DNA sequences used in flag knockins.

Donor DNA	Sequence
N-terminal 3xflag-CHD4	GGGTGGGGGCAGGCTGTTGTTCAAAAGTGCATCCATATCCTCCTCCTCAC TGCCCGCCGAGCAGGGGGACGGGGAGCCCAGGCCCGACGCTTTATCGTCA TCATCTTTGTAGTCCTTGTCATCATCGTCCTTATAGTCCTTATCGTCGTC ATCCTTGTAATCCATCCCCTTCCGCTCCCGGCCAGGGAATTGGCCCAGCT
C-terminal 3xflag-CHD4	TATCAGTGAATTCCCTTGAGATTGGGTGGATTCCTATCATCTGGAATTCT GACTCCTCGGGTCTCCTTCTCTAGGTAGCACAGCAACAGGATTACAAGGA TGACGACGATAAGGACTATAAGGACGATGATGACAAGGACTACAAAGATG ATGACGATAAATAGAGCAGTGAAGATGCAGACTGATACCACCTCCACCGC
N-terminal 3xflag-BRD4	TTGGGCCTGGGCCTGTGTTGTAGACATTTGGGAAGTTTCTAGTCCATCCC CCATTACTGGCAGATTTCTCAATCTCGTCCCGGGGCCGCTCTCCGCAGAC TTGTCGTCATCGTCTTTGTAGTCCTTGTCGTCATCGTCTTTGTAGTCCTT GTCGTCATCGTCTTTGTAGTCCATGCTAGTGATCCCATCACATTCTTCAC

Supplementary Table 5 Antibodies used in ChIP assays.

Antibody target	Catalog Number	Company
Flag	F1804	Sigma Aldrich
HDAC2	Ab7029	Abcam
RBBP4	A-2703-050	Epigentek
BRD4	A301-985A100	Bethyl Laboratories
H3K9ac	61663	Active Motif
H3K9me1	39887	Active Motif
H3K9me3	39765	Active Motif
RNA Pol 2	61667	Active Motif

ChIP-seq data for H3K27ac, H3K27me3, H3K36me3, H3K4me1, H3K4me2, H3K4me3, BRD4, MYOD , MYOG, MYCN, CTCF, Rad21 in RH4 cells were obtained from[15].

Supplementary Table 6 Sequence of guide RNAs used in NuRD-centered CRISPR screen.

Target	Sequence
CHD4	CTCCTCACTGCCCGCCGAGCAGG
	TTAGGGATTTTAGGGTCCCGAGG
	AAGAGGGGCGCTCACTAGGAAGG
	ATGTATGTCGTAACCTATGTGGG

	TATAAATACGAACGCATCGATGG
CHD3	ACGAGGGAGTACTCGGGCGCGGG
	GGAGTACTGCCGCGTATGCAAGG
	CTCGCCGGAGCCATAGTCCAGGG
	GAATTCATTCTCACGAATGATGG
	GCGCATCGATGGTGGTATCACGG
HDAC1	CGCAGACGCAGGGCACCCGGAGG
	TGAGTCATGCGGATTCGGTGAGG
	TTACGTCAATGATATCGTCTTGG
	TATTCACCATGGTGACGGCGTGG
	CATCAATCCCGTCTCGGAGCGGG
HDAC2	TGGGTCATGCGGATTCTATGAGG
	TACAACAGATCGTGTAATGACGG
	GATGTATCAACCTAGTGCTGTGG
	CTCATTATCTGGTGATAGACTGG
	TCCGTAATGTTGCTCGATGTTGG
LSD1	CGGACCCCGGCGGTTCCGCCAGG
	GTCGGACCAGCCGGCGCAAGCGG
	TGAGAAGTCATCCGGTCATGAGG
	CGAGTTGCCACATTTCGCAAAGG
	GAACTCGGCAGTAATATCTCTGG
MBD2	ATGCGCGCGCACCCGGGGGGAGG
	CGGCGACTCCGCCATAGAGCAGG
	CGCCACTCGGGGGACGGCCGCGG
	CCTTCTTCCATCCGGGGGGGAGG
	GCCTCAGTTGGCAAGGTACCTGG
MBD3	GAGTGCCCGGCGCTCCCGCAGGG
	CTTGCCCGTGCGGAAGTCGAAGG
	GCTTGAAGATGGACGCCGTCTGG
	CAGCAATGTCGAAGGCGTTCAGG
	CGTGGTGTTGAGCCATACGCCGG
MTA1	CGATTCTCCGGATCAGGTATGGG
	CTGCTATAAGGCCGGACCGGGGG
	CGACTCGGTCTCGTTGAGCAGGG
	TCGAGTAGGAAACCGGTACCAGG
	ATCAGAGGCCAACCTTTTCGAGG
MTA2	TTACCTGGTTAGACGGATTGAGG
	TGTCTGTCTTTTCCGGCGCAGGG
	GATCCCAGATCGCCTAGTAGAGG
	GCAAAGGAACGGCTACGACCTGG
	GGACTTCAATGATATTCGCCAGG
MTA3	GCAACCCATACCTAATAAGAAGG
	TGCGAGCATTATAAGTGTGTTGG
	CTCTGCCCGCAACACATATCAGG
	TCAATCTGTCGATCCGTAAGTGG
	AGTGTCTTAGTACCACTCGGAGG
GATAD2A	CCGAAGAAGCATGCCGAACACGG

	GCGGCAGAGTCAAATACAAAAGG
	GCGGCCGCCTGTCGGCTTGCTGG
	AACTCGTTGTTGGCGGCGCTGGG
	AGACGGACTTCACGTGCCGCTGG
GATAD2B	GAATCTGTTGAAGCGGAGCTTGG
	ACGATTGGAAGAAGCCCGACTGG
	GGCAGCATCAGTCATGGCGCTGG
	CATGTAGATGAACTCGCTATTGG
	GGTGAAATCTGTGCGGCACTGGG
RBBP4	TTCTTCCACTGCGTCGTCGAAGG
	GTCTTTGTTGCGATGATACAAGG
	AGACTCATGGAGTAGATGCCAGG
	CAGTGTGAGCATCAACTGAGTGG
	CAGTCTGCGATCAGTACCACTGG
RBBP7	TACACCGTTTCTATATGACCTGG
	GCCCTTCATTGGCTAGTGCTGGG
	GGTGGCTTTGGTTCTGTAACAGG
	TGTCTGTGGGATATAAACGCAGG
	CTGCTGCACGAGTCATTGTTTGG
BRD4	AGATTTCTCAATCTCGTCCCAGG
	TAAGATCATTAAAACGCCTATGG
	TGATGTCACAGTAGTCGTGTAGG
	CCTGAGCATCACGGTACTCACGG
	AGTCGATTTCAATCTCGTCGGGG

Supplementary Table 7 Primers used in ChIP-qPCR

Primer Name	Sequence
ALK_FW	GTCACTTTGGGTCACTTGCT
ALK_RV	GCCTTGTAGTTAGCTCTCCC
ASS1_FW	CAATGGTGGAGCGTGAAAT
ASS1_RV	ACCCTCCCATTCTCTTTGC
CDH3_FW	ATGCTCCCGAGATACCAGAT
CDH3_RV	AGAAGCGTTGTAATCCTCCAA
UNTR5_FW	TATAAAGGACCGTGGCTTCC
UNTR5_RV	TCATTCATTTGGTCATGGCT

Supplementary Table 8 TaqMan gene expression assays.

Gene name	Assay catalogue number
ALK	Hs00608284_m1
ASS1	Hs01597981_gH
CDH3	Hs01285856_cn
CHD4	Hs00172349_m1
CNR1	Hs01038522_s1
GAPDH	Hs02758991_g1
PAX3-FOXO1	Hs03024825_ft
PIPOX	Hs04188864_m1
RBBP4	Hs01568507_g1
TFAP2B	Hs00231468_m1

Western Blot

Cells were harvested and lysed in standard lysis buffer (50mM Tris-HCl pH 7.5, 150mM NaCl,1% NP-40, 0.5% sodium deoxycholate, 0.1% SDS, 1mM EGTA, 50mM NaF, 5mM $Na_4P_2O_7$, 1mM Na_3VO_4, and 10mM ß-glycerolphosphate with protease inhibitor cocktail, cOmplete Mini®, Roche). Lysates were run through 20G needles and cleared by centrifugation. Protein concentration was measured with Pierce™ BCA Protein Assay Kit (#23227, Thermo Scientific) according with the manufacturer's instructions. Proteins were separated with NuPAGE 4-12% Bis-Tris pre-cast gels (ThermoFisher Scientific AG) and transferred into nitrocellulose membranes (GE Healthcare). Membranes were blocked with 5% milk in TBST and primary antibodies (1:1000) were incubated over night at 4°C. After TBST wash, the membranes were incubated for 1hr with HRP-linked secondary antibodies at room temperature. Proteins were detected by chemiluminescence using Amersham™ ECL™ detection reagents (#RPN2209, GE Healthcare) or SuperSignal™ West Femto Maximum Sensitivity Substrate (#34095, ThermoFisher Scientific AG). The following commercial primary antibodies were used: anti-CHD4 (#A301-082A, Bethyl Laboratories), anti-Flag (clone M2, #F1804 , Sigma Aldrich), anti-BRD4 (#ab128874, Abcam), anti-FOXO1 (H-128, #sc-11350, Santa Cruz Biotechnology), anti-HDAC2 (3F3, #5113S, Cell Signaling Technologies), anti-MTA2 (#M7569, Sigma Aldrich), anti-RBBP4 (A301-206A-M, Bethyl Laboratories) and anti-βTubulin (D3U1W, #86298S, Cell Signaling Technologies), and anti-GAPDH (14C10, #2118L, Cell Signaling Technologies).

Immunofluorescence

Cells were seeded on cover slides, fixed with 4% PFA for 15 min, followed by a 5 min incubation with 0.1M glycine in PBS, permeabilized for 15 min with 0.1% Triton X-100 in PBS and blocked with 4% horse serum in 0.1% Triton X-100/PBS. All steps were carried out at room temperature. Then, cells were incubated overnight with anti-flag antibody (clone M2, #F1804, Sigma Aldrich) at 1:250 dilution in 4% horse serum in 0.1% Triton X-100/PBS. Fluorescent secondary antibody (Alexa Fluor™ 594 anti-mouse, #A11032, Thermo Fisher Scientific AG) at 1:200 dilution in PBS containing 4% horse serum was applied for 1hr at room temperature. Cover slides were fixed on objective glass with DAPI Vectashield® mounting medium (Vector laboratories Inc.) and analyzed by fluorescence microscopy.

Cell count assay

RH4 cells with or without flag knockin were platted (45 000 cells/well in 24-well plate) in duplicate and counted every day for 6 days. Three biological replicates were performed.

ChIP-qPCR

ChIP assays were performed using the same protocol as for ChIP-seq. Quantitative PCR was performed using PowerUp™ SYBR Green Master Mix (ThermoFisher Scientific AG) and primers were designed to target known binding sites of PAX3-FOXO1, HDAC2 and BRD4. As negative control, we used an untranscribed genomic region (UNTR5) and the commercially available Human Negative Control Primer Set 1 (#71001, Active Motif) was used for normalization. Primer sequences used can be found in the Supplementary Table 7.

Quantitative real time PCR

Total RNA was extracted using the RNeasy mini Kit (#74106, Qiagen Instruments AG). cDNA synthesis was carried out using the High-Capacity Reverse Transcription Kit (#4368814, Applied Biosystems by ThermoFisher Scientific AG). Quantitative PCR was performed using TaqMan gene expression master mix and TaqMan gene expression assays (Applied Biosystems by ThermoFisher Scientific AG). Data was analyzed with SDS 2.3 software and Ct values were normalized to GAPDH. Relative expression levels were calculated using the ΔΔCt method based on experiments performed in triplicates. Details regarding the gene expression assays used can be found in Supplementary Table 8.

Library preparation for ChIP-seq assays

DNA libraries were created using Illumina TruSeq ChIP Library Prep Kit, after which DNA was size selected with SPRIselect reagent kit (to obtain a 250-300 bp average insert fragment size). Libraries were multiplexed and sequenced using NextSeq500 High Output Kit v2 (75 cycles), #FC-404-2005 on an Illumina NextSeq500 machine. 25,000,000-30,000,000 unique reads were generated per sample.

ChIP-seq data processing

ChIP enriched DNA reads were mapped to reference genome (version hg19) using BWA[1]. Duplicate reads were discarded. For IGV sample track visualization, coverage density maps (tdf files) were generated by extending reads to the average size (measured by Agilent Bioanalyzer minus 121 bp for sequencing adapters) and counting the number of reads mapped to each 25bp window using igvtools (https://www.broadinstitute.org/igv/igvtools). ChIP-seq read density values were normalized per million mapped reads.

ChIP-seq Peak calling

High-confidence ChIP-seq peaks were called by MACS2 (https://github.com/taoliu/MACS) with the narrow algorithm[2]. The peaks which overlapped with the possible anomalous artifact regions (such as high-mappability regions or satellite repeats) blacklisted by the ENCODE consortium (https://sites.google.com/site/anshulkundaje/projects/blacklists) were removed using BEDTools. Peaks from ChIP-seq were selected at a stringent p-value of at least 0.0000001. Peaks within 2,500bp to the nearest TSS were set as promoter proximal, while all other were considered distal. The distribution of peaks (as intronic, intergenic, exonic, etc.) was annotated using HOMER.

Chromatin state mapping

Chromatin states were identified using a hidden Markov model to find complex patterns of chromatin modifications (http://compbio.mit.edu/ChromHMM/) which was developed for the ENCODE project[3,4]. We defined chromatin states by integrated analysis of 9 histone modifications (H3K27ac, H3K27me3, H3K4me1, H3K4me2, H3K4me3, H3K36me3, H3K9ac, H3K9me1, and H3K9me3) and 2 architectural proteins (RAD21 and CTCF), allowing us to identify 16 states in RH4 cells.

Motif enrichment analysis

Enrichment of known motifs was performed using HOMER script "findMotifsGenome.pl" (http://homer.salk.edu/homer/ngs/peakMotifs.html).

Visualization of ChIP-seq and DNase-HS data

Metagene plots and heatmaps of ChIP-seq and DNAse-HS data were performed using NGSplot[5]. Scaling and plot adjustments were performed using replot.r function (https://github.com/shenlab-sinai/ngsplot/wiki/ProgramArguments101). Colors for heatmaps were set in Adobe Photoshop, while sizing and placement were performed in Adobe Illustrator.

RNA-seq library preparation and data analysis

Library preparation and sequencing was performed by ATLAS Biolabs GmbH, Berlin, Germany. Poly-A selected cDNA libraries were prepared with the Illumina® TruSeq® Stranded mRNA Sample Preparation Kits and sequenced on an Illumina HiSeq2000. RNA-seq reads were aligned to the Ensembl release 91's GRCh38 reference genome with the options "--rna-strandness RF --fr"(ftp://ftp.ensembl.org/pub/release91/fasta/homo_sapiens/dna/Homo_sapiens.GRCh38.dna_sm.primary_assembly.fa.gz) using Hisat2 v2.1.0[6] and Samtools v1.7[7]. Mapping quality was assessed by qualimap v2.2.1[8].

Transcripts from mapped reads were assembled using the Ensembl release 91's reference transcriptome(ftp://ftp.ensembl.org/pub/release91/gtf/homo_sapiens/Homo_sapiens.GRCh38.91.gtf.gz) and StringTie v1.3.4[9]. Gene expression values were calculated at the gene level from the assembled transcripts using featureCounts v1.6.0 with the options of "reverse strand-specific" and "pairs with both ends mapped"[10] based on the merged transcriptome of all samples generated by StringTie[11].

Differential gene expression was analyzed using the package DESeq2 v3.7[12] (http://bioconductor.org/packages/devel/bioc/vignettes/DESeq2/inst/doc/DESeq2.html) of R v3.4.3 with the false discovery rate set at 0.01 and p-value at 0.05. PCA analysis, clustering of gene expression by the Pearson correlation method, and clustered heatmaps using the pheatmap package (https://www.rdocumentation.org/packages/pheatmap) were all carried out in R v3.4.3.

Gene Set Enrichment Analysis

Enrichment for curated gene sets (Molecular Signatures Database v6.2) was performed using GSEA software version 3.0[13,14]. The GSEA Preranked tool was employed with 1000

permutations, and only gene sets with a maximum list size of 500 were considered. The differentially expressed gene lists were pre-ranked by the metric value calculated as log10(pvalue) and divided by the reverse sign of log2(fold-change).

SUPPLEMENTARY REFERENCES

1. Langmead, B., Trapnell, C., Pop, M. & Salzberg, S. L. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* **10,** R25 (2009).
2. Zhang, Y. *et al.* Model-based Analysis of ChIP-Seq (MACS). *Genome Biol.* **9,** R137 (2008).
3. Ernst, J. *et al.* Mapping and analysis of chromatin state dynamics in nine human cell types. *Nature* **473,** 43–49 (2011).
4. Ernst, J. & Kellis, M. ChromHMM: automating chromatin-state discovery and characterization. *Nat. Methods* **9,** 215–216 (2012).
5. Shen, L., Shao, N., Liu, X. & Nestler, E. ngs.plot: Quick mining and visualization of next-generation sequencing data by integrating genomic databases. *BMC Genomics* **15,** 284 (2014).
6. Kim, D., Langmead, B. & Salzberg, S. L. HISAT: a fast spliced aligner with low memory requirements. *Nat. Methods* **12,** 357–360 (2015).
7. Li, H. *et al.* The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25,** 2078–2079 (2009).
8. García-Alcalde, F. *et al.* Qualimap: evaluating next-generation sequencing alignment data. *Bioinformatics* **28,** 2678–2679 (2012).
9. Pertea, M. *et al.* StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. *Nat. Biotechnol.* **33,** 290–295 (2015).
10. Liao, Y., Smyth, G. K. & Shi, W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* **30,** 923–930 (2014).
11. Pertea, M., Kim, D., Pertea, G. M., Leek, J. T. & Salzberg, S. L. Transcript-level expression analysis of RNA-seq experiments with HISAT, StringTie and Ballgown. *Nat. Protoc.* **11,** 1650–1667 (2016).
12. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15,** 550 (2014).
13. Subramanian, A. *et al.* Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci. U. S. A.* **102,** 15545–50 (2005).

14. Mootha, V. K. *et al.* PGC-1α-responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes. *Nat. Genet.* **34,** 267–273 (2003).
15. Gryder, B. E. *et al.* PAX3-FOXO1 Establishes Myogenic Super Enhancers and Confers BET Bromodomain Vulnerability. *Cancer Discov.* **7,** 884–899 (2017).

SUPPLEMENTARY FIGURES

Supplementary Figure 1

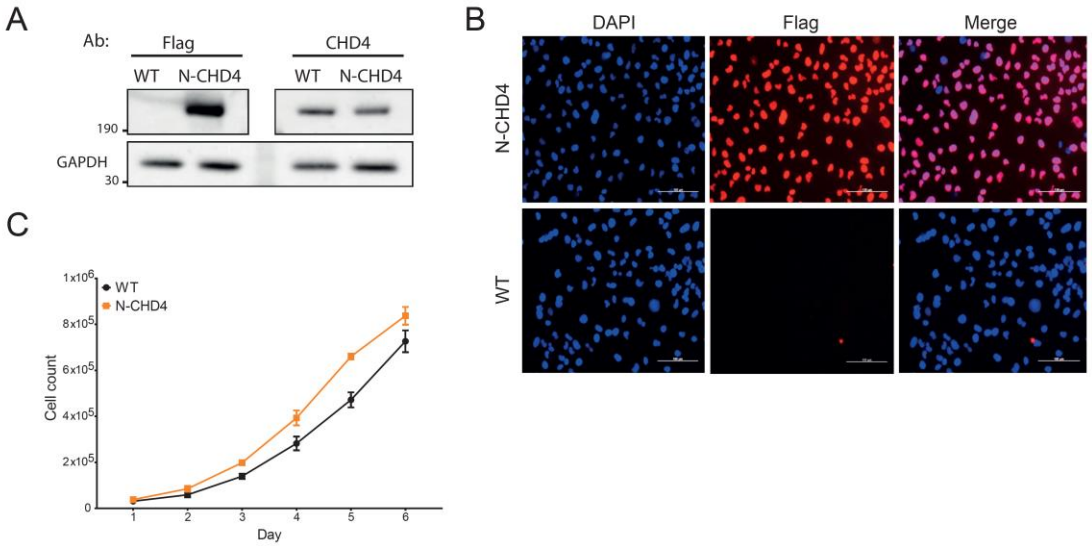

Fig.S1 – CRISPR/Cas9 efficiently adds a 3xflag tag in-frame to the N-terminal of endogenous CHD4 in RH4 cells

(A) Immunoblot, performed with anti-flag and anti-CHD4 antibodies, confirms the insertion of the 3xflag tag in the N-terminal of CHD4 in RH4 cells (named N-CHD4). GAPDH was used as a loading control. Wildtype RH4 cells (WT) served as negative control. (B) Immunofluorescence, performed with an anti-flag antibody, shows the expected nuclear localization of the flag tagged CHD4 in N-CHD4 cells. DAPI was used to visualize the nucleus and wildtype RH4 cells (WT) as negative control. Scale bar - 100µm. (C) Cell counts throughout 6 days of wildtype RH4 cells (WT) and N-CHD4 cells.

Supplementary Figure 2

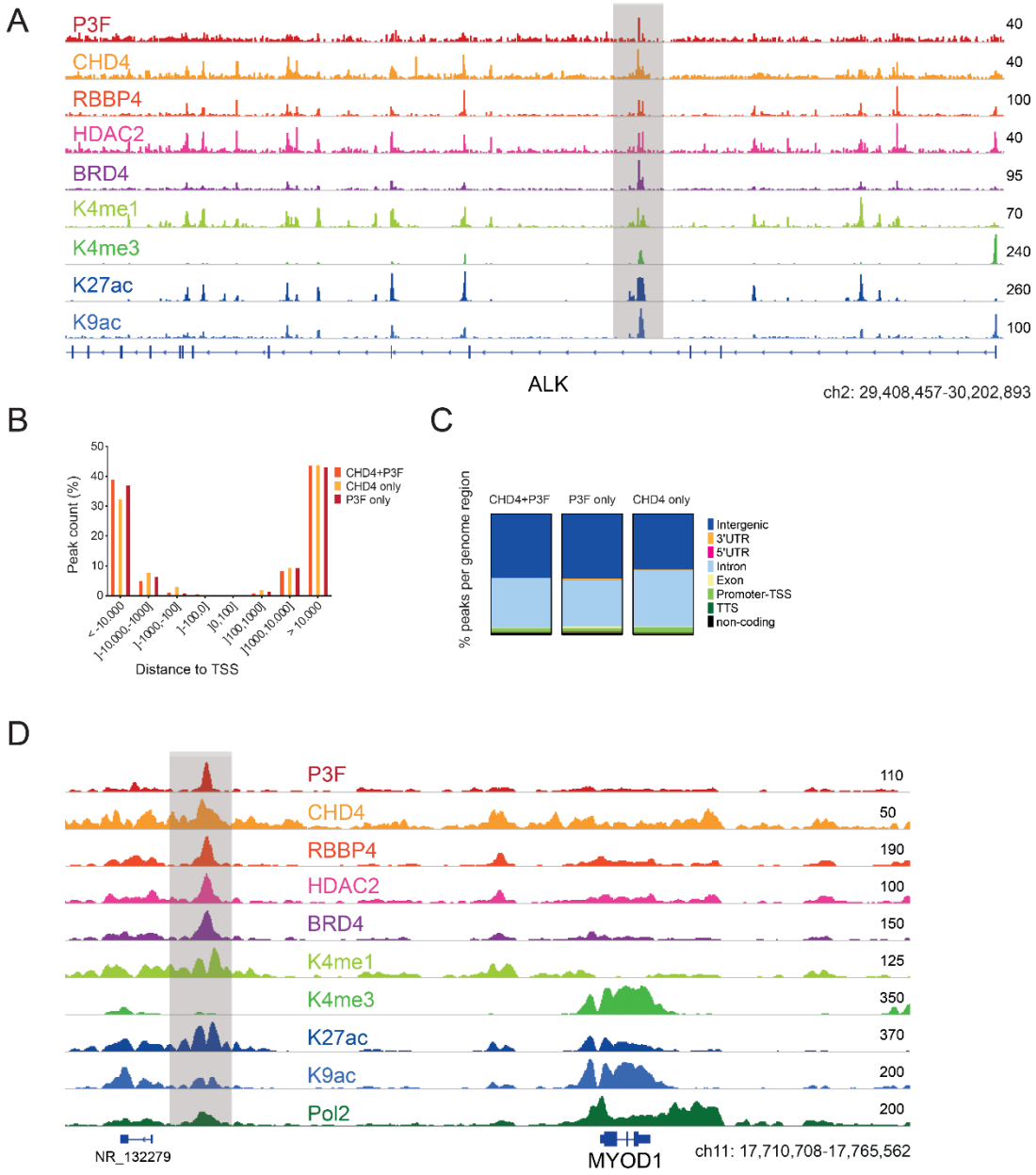

Fig.S2 – CHD4 binds to distal intronic and intergenic enhancer regions
(A) Example of gene tracks of the indicated proteins at the intronic enhancer present in the ALK gene. (B) Distribution of the peak counts according to the distance to the transcription start site (TSS), in percentage, of CHD4 or P3F within CHD4+P3F, CHD4-only and P3F-only regions (C) Distribution of the peak counts according to their genome location, in percentage, of CHD4 or P3F within CHD4+P3F, P3F-only, and CHD4-only regions. (D) Example of gene tracks of the indicated proteins at the MYOD1 (myogenic differentiation 1) SE loci in RH4 cells.

Supplementary Figure 3

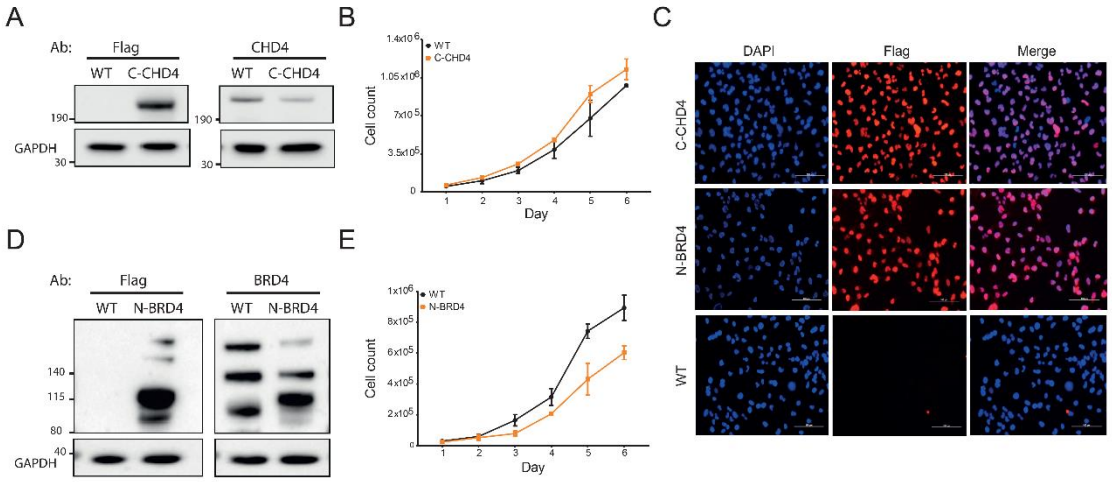

Fig.S3 – CRISPR/Cas9 efficiently adds a 3xflag tag in-frame to the C-terminal of endogenous CHD4 protein and to the N-terminal of endogenous BRD4 protein in RH4 cells
(A) Immunoblots, performed with anti-flag and anti-CHD4 antibodies, confirm the insertion of the 3xflag tag in the C-terminal of CHD4 in RH4 cells (named C-CHD4). GAPDH was used as a loading control and wildtype RH4 cells (WT) as negative control. (B) Cell counts throughout 6 days of wildtype RH4 (WT) and C-CHD4 cells. (C) Immunofluorescence, performed with an anti-flag antibody, shows the expected nuclear localization of flagged CHD4 in C-CHD4 cells and of flagged BRD4 in N-BRD4 cells. DAPI was used to visualize the nucleus and wildtype RH4 cells (WT) served as negative control. Scale bar - 100µm. (D) Immunoblots, performed with anti-flag and anti-BRD4 antibodies, confirm the insertion of the 3xflag tag in the N-terminal of BRD4 in RH4 cells (named N-BRD4). (E) Cell counts throughout 6 days of wildtype RH4 cells (WT) and N-BRD4 cells.

Supplementary Figure 4

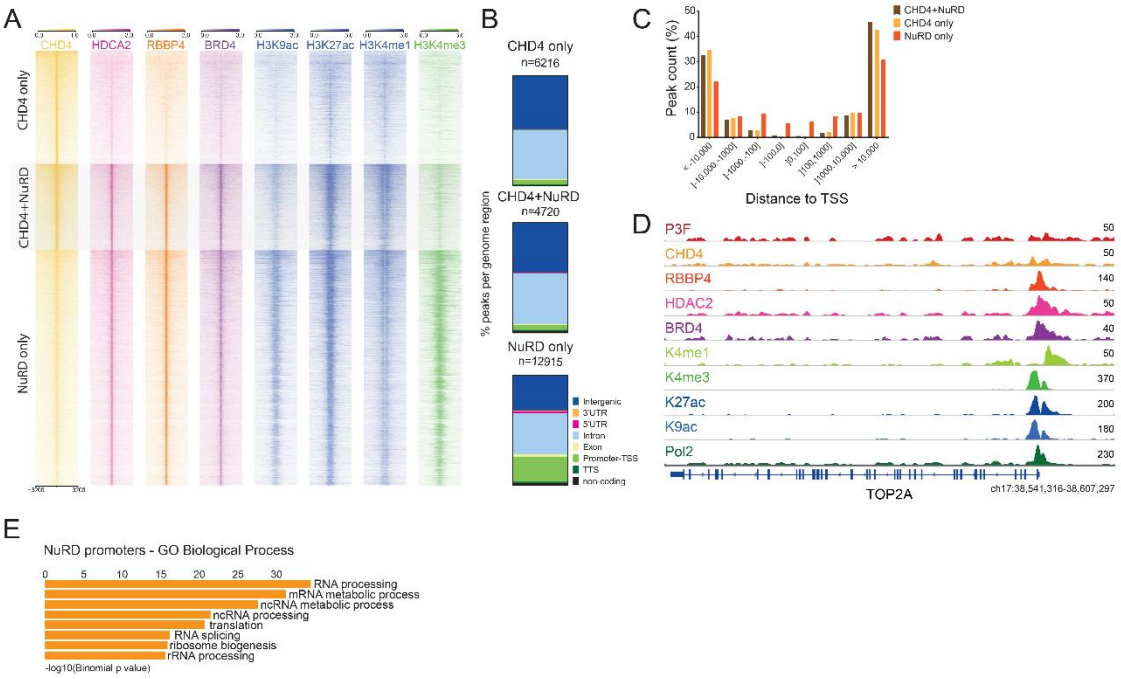

Fig.S4 – CHD4-free NuRD localizes to promoters together with BRD4
(A) Heatmaps depicting the signal of the indicated proteins at CHD4-only (n=6216), CHD4+NuRD (n=4720) and NuRD-only regions (n=12915) in the genome of RH4 cells. NuRD corresponds to the overlap between HDAC2 and RBBP4 ChIP signal. The rows show 7kb regions, centered on CHD4 or HDAC2 ChIP peaks and ranked by the signal intensity of H3K27ac. Color shading corresponds to ChIP-seq read counts. (B) Distribution, in percentage, of the peak counts of CHD4 or NuRD within CHD4-only, CHD4+NuRD and NuRD-only regions according to their genome location. (C) Distribution of the peak counts, in percentage, of CHD4 or NuRD within CHD4+NuRD, CHD4-only or NuRD-only regions according to the distance to the transcription start site (TSS). (D) Example of gene tracks of the indicated proteins at the CHD4-free NuRD-containing TOP2A (DNA topoisomerase II alpha) promoter in RH4 cells. (E) Gene ontology analysis (assigned by GREAT) of the biological processes associated with the promoters bound by NuRD-only.

Supplementary Figure 5

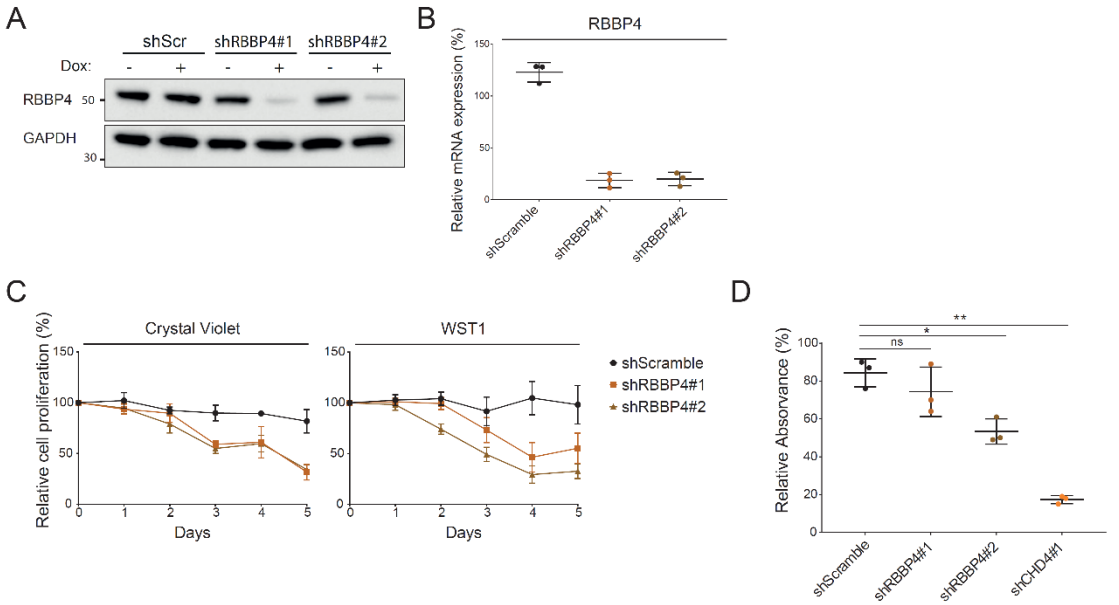

Fig.S5 – RBBP4 silencing influences FP-RMS cell proliferation
(A) Immunoblotting confirms the knockdown (KD) of RBBP4 achieved with 2 doxycycline-inducible shRNAs (shRBBP4#1/2) in RH4 cells after 72hrs of treatment (Dox). RH4 cells expressing a scrambled shRNA (shScr) served as negative control and GAPDH was used as loading control. (B) Expression levels of RBBP4 quantified by qPCR in RH4 cells after 72hrs of doxycycline treatment. Expression levels (relative to GAPDH) were normalized to uninduced cells. Data represented as mean ± SD (n=3). (C) RH4 cell proliferation was measured by crystal violet and WST1 assays at the indicated time points after silencing of RBBP4 by shRNA induction with doxycycline. Data is represented as mean ± SD (n=3) and was normalized to uninduced cells. A shScramble construct served as negative control. (D) Cell proliferation measured by BrdU incorporation in RH4 cells after 72hrs of RBBP4 or CHD4 KD. Data is represented as percentage of absorbance at 450nm normalized to uninduced control. A scrambled shRNA was used as negative control. (* p< 0.1, **p < 0.01; ***p < 0.001, ratio paired t test).

Supplementary Figure 6

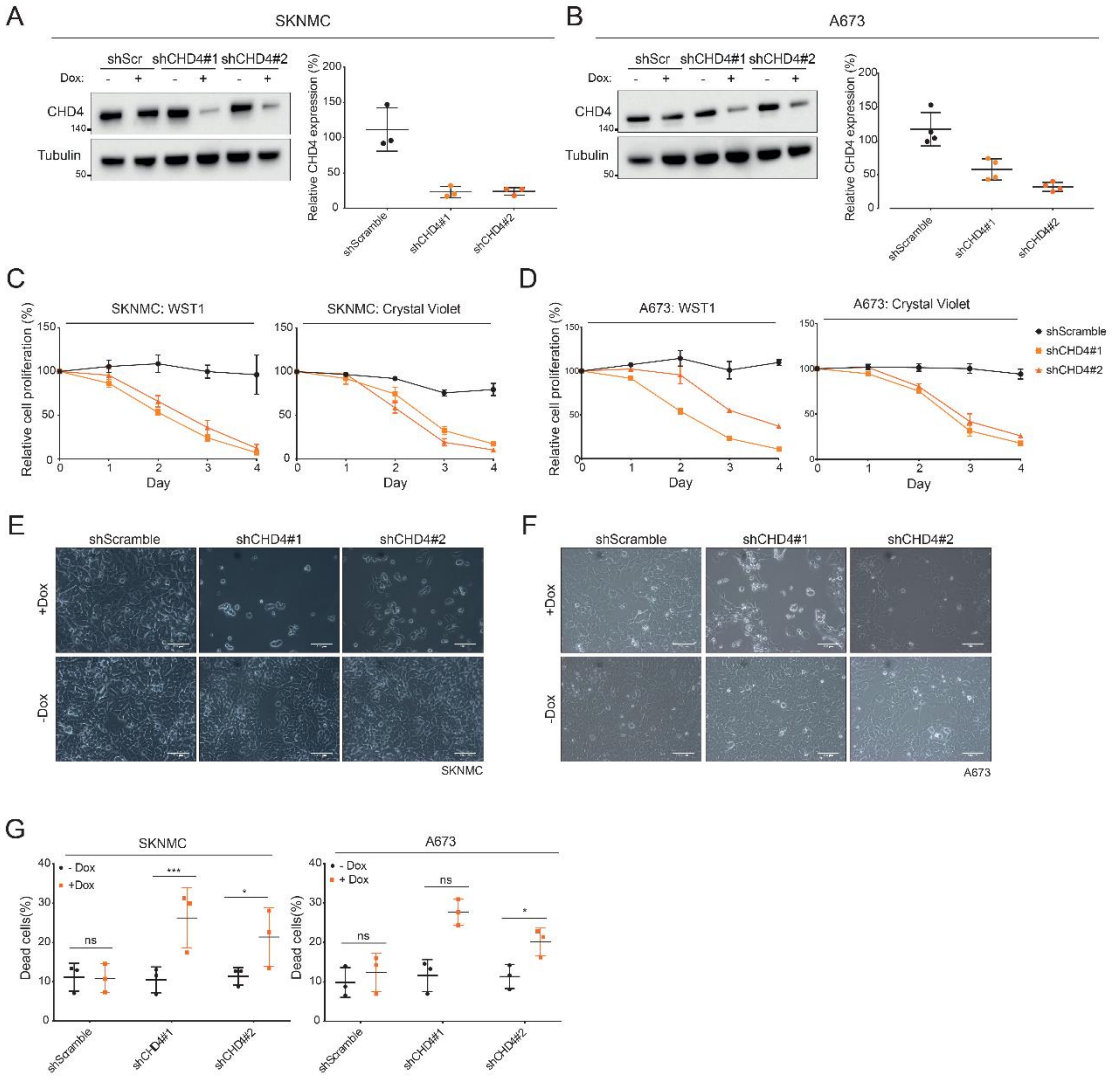

Fig.S6 – CHD4 silencing induces Ewing sarcoma cell death
(A/B) *Left panels* Western blot validates the knockdown obtained in SKNMC and A673 cell lines after 48hrs of doxycycline-induction of 2 shRNAs targeting CHD4. Tubulin was used as loading control. *Right panels* Expression levels of CHD4, quantified by qPCR, in SKNMC and A673 cells after 72hrs of silencing induction. Expression levels (relative to GAPDH) were normalized to uninduced cells. Data represented as mean ± SD (n of at least 3). (C/D) Cell proliferation of SKNMC and A673 cell lines measured by crystal violet and WST1 assays at the indicated time points after silencing of CHD4 by 2 doxycycline-inducible shRNAs. Data is represented as mean ± SD (n=3) and was normalized to uninduced cells. A shScramble construct served as negative control. (E/F) Representative phase-contrast images of SKNMC and A673 cells 72hrs after doxycycline-induction of 2 shRNAs targeting CHD4. Scale bar - 100µm. (G) Percentage of dead cells observed 72hrs after induction of CHD4 silencing by doxycycline treatment. Cells were stained with 7AAD and analyzed by FACS. A scrambled shRNA served as negative control. Values represented as mean ± SD of 3 independent experiments (*p< 0.1, **p < 0.01; ***p < 0.001, ratio paired t test).

Manuscript 2

Structure-based drug design of small molecule inhibitors targeting the PHD2 reader domain of CHD4

Joana G. Marques[1†], Dimitrios Spiliotopoulos[2†], Zuzana Becarova-Zigova[3], César Fernández[4], Amedeo Caflisch[2], Beat W. Schäfer[1]

[1]Department of Oncology and Children's Research Center, University Children's Hospital Zurich, Zurich, Switzerland
[2]Department of Biochemistry, University of Zurich, Zurich, Switzerland
[3]Institute of Molecular Biology & Biophysics, Department of Biology, Swiss Federal Institute of Technology in Zurich – ETH, Zurich, Switzerland
[4]Novartis Institutes for BioMedical Research (NIBR), Basel, Switzerland

[†]These authors contributed equally to the presented study.
Conflict of Interest: The authors declare no conflict of interest.

Manuscript in preparation.

Corresponding Author: Beat W. Schäfer

Department of Oncology, Children's Hospital Zurich
Steinwiesstrasse 75
8032 Zurich, Switzerland
Beat.schaefer@kispi.uzh.ch
Phone: +41 (44) 266 7553
Fax: +41 (44) 634 8859

Author Contributions: JGM and ZBZ purified and labelled the recombinant proteins. DS performed the *in silico* analysis. JGM performed all the NMR assays. JGM, DS, CF, AC and BWS contributed to the analysis and interpretation of data and made substantial contributions to conception, design, and intellectual content of the paper. JGM, DS and BWS wrote the paper.

ABSTRACT

CHD4 is a chromatin remodeler involved in transcription regulation, DNA-damage repair, and tumor development. In cancer, overexpression of CHD4 is associated with poor prognosis and its silencing causes cell death in a variety of tumors, such as breast cancer, acute myeloid leukemia, and fusion-positive rhabdomyosarcoma. Nevertheless, no inhibitors targeting this protein have been developed yet.

CHD4 is composed of a C-terminal ATPase, 2 plant homeodomains (PHD1/2), 2 chromodomains and 2 C-terminal domains of unknown function (DUF). Since high resolution structural information is available for the PHD domains and the ATPase activity is dependent on the binding of these domains to histones, we aimed to target this chromatin remodeler by developing small molecule inhibitors against its second PHD domain. We hypothesized that by preventing the binding of CHD4 to histones we might be able to interfere with the remodeling activity of CHD4 and hence affect tumor development.

By structure-based drug design we originated a comprehensive library of potential PHD2 binding small molecule inhibitors. These candidates were tested for binding to recombinant PHD2 by NMR, both by ligand and protein observation assays. From the 105 compounds tested, we identified one weak binder. To our knowledge, this is the first attempt on targeting a H3K9-binding PHD domain. Although the PHD ligand reported here should undergo optimization cycles to improve its binding affinity, this study shows that the PHD reader domain, otherwise deemed undruggable, can still be considered as a feasible target for structure-based drug design efforts.

Significance

- First report of a structure-based drug design study targeting a H3K9-binding PHD domain and targeting the chromatin remodeler CHD4
- Identification of weak small molecule binding the second PHD domain of CHD4

Keywords

Plant homeodomain, NMR, SEED, Pharmit, CHD4

INTRODUCTION

Nucleosomes are the fundamental packaging units of our genome and they allow the fitting of nearly 2 meters of DNA into a 10-20µm nucleus[1]. These packaging units are formed by two copies of four histone proteins, H2A, H2B, H3, and H4, that assemble into an octamer, around which 145–147 base pairs of DNA are wrapped[2].

Histones are subjected to many post-translational modifications (PTMs), such as methylation and acetylation, that influence DNA compaction and hence the accessibility of the genome to the transcription machinery[3]. These PTMs, commonly named histone marks, work as docking sites of effector proteins that possess catalytic activity and can influence chromatin structure. The docking of these effectors is mediated by "reader" domains that recognize and bind to specific histone marks. For example, acetylation is recognized by bromodomains, while methylation is recognized by chromo, tudor, MBT (malignant brain tumor), and PHD domains[4]. These epigenetic "readers" are promising therapeutic targets as shown by the success story of JQ1, the small molecule inhibitor of the acetyl-lysine binding BET-bromodomain of BRD4[5,6]. Since its development, many studies have shown in vitro and in vivo benefits of JQ1 treatment in many hematological malignancies as well as in a variety of solid tumors such as glioblastoma, medulloblastoma, hepatocellular carcinoma, colon cancer, pancreatic cancer, prostate cancer, lung cancer, and breast cancer[7].

Plant homeodomains (PHD) are small (circa 65 residues) zinc-finger "reader" domains found in many chromatin-remodeling proteins. Mutations in these domains are linked to several human diseases, including cancer, mental retardation, and immunodeficiencies. PHD fingers can be subdivided into groups according to the histone mark that they recognize, such as H3K4me3, H3K9me3, H3/H4 acetylated at various lysines, or unmodified tail of histone 3[8].

CHD4 (chromodomain-helicase-DNA binding protein 4) is an ATPase from the SNF2-like family with nucleosome remodeling activity[9,10]. This chromatin remodeler collaborates with oncogenic transcription factors, like PAX3-FOXO1 and EWS-FLI1, and coregulates aberrant gene expression[11,12]. Moreover, CHD4 silencing causes tumor cell death in fusion-positive rhabdomyosarcoma, breast cancer, and acute myeloid leukemia[11,13,14], making it an attractive target for therapy. This remodeler is composed of a C-terminal ATPase, 2 plant homeodomains (PHD1/2), 2 chromodomains and 2 C-terminal domains of unknown function (DUF)[9]. Its chromodomains bind directly to DNA, while its PHD domains interact with histones[15]. These PHD domains recognize the N-tail of histone 3, with increased affinity when Lys9 is methylated or acetylated[16–18] and in the absence of methylation of Lys4 or acetylation of Ala1[18]. Importantly, the ATPase activity of CHD4 is dependent on the simultaneous binding of the PHD domains to histones and of the chromodomains to

DNA[19,20]. Hence, we hypothesized that, by preventing the binding of the PHD domains to histone 3, we could interfere with the remodeling activity of CHD4.

In this study, we aimed to develop the first in-class small molecule inhibitor of CHD4 by targeting its second PHD domain. Structure-based drug design allowed us to compose a library of potential PHD2-CHD4 small molecule inhibitors which were validated by NMR protein and ligand observation experiments.

RESULTS

The structure of PHD2-H3 interaction informs *in silico* screens

The second PHD finger of CHD4 is a "reader" of the H3 N-terminal tail. This zinc finger has a very loose conformation, comprised of a small β-sheet and an α-helix, which are held together by the coordination with two zinc ions (Fig.S1A). PHD2 recognizes unmodified H3, H3K9me1/2/3 and H3K9ac[18]. The H3 tail occupies an elongated binding site on PHD2, forming a third antiparallel β-strand which pairs with the first β-strand of PHD2, and is held in place through a conserved set of salt bridges and hydrogen bonds (Fig.S1B). The binding is stabilized by backbone interactions between Arg2, Lys4, and Thr6 of H3 with the residues Cys513, Leu511, Gly500 or Gly508. Further, the interaction is restrained by the co-ordination of the amino group of Ala1 by Gly536 and Glu537, the formation of a salt bridge between Arg2 and Asp515, and hydrogen bonding interactions involving Lys4, Thr6 and Arg8. Methylation or acetylation of Lys9 increases the affinity of the binding, since it allows Arg8 to form a salt bridge with Glu500. Moreover, the aromatic ring of a phenylalanine lies orthogonal to the protein surface and may favorably interact with H3K9ac or form hydrophobic and cation-π contacts with methylated Lys9.

Experimental evidence[21,22] indicates that the primary sequence and post-translational modification status of the first residues of the H3 N-terminal histone tail is crucial to the binding to the PHD finger, whereas computational results suggest that the latter residues (mainly, H3R8 and H3K9) modulate the fine tuning of binding[23]. Since we aimed to abolish the PHD finger/histone tail binding, we decided to focus our computational efforts on the PHD finger surface establishing interactions with residues H3A1–H3K4 (Fig.1A). In particular, we undertook two different *in silico* screening approaches: two screening campaigns were performed using SEED (Fig.1B/C)[24,25] and one was performed with Pharmit (Fig.1D)[26]. From these three *in silico* screens, a library of 105 candidate small molecules binding the PHD2 domain of CHD4 was generated and purchased. More details regarding these screens are available in the material and methods section.

The purified PHD2-CHD4 domain is functional

To test the candidate compounds, we purified the PHD2 domain of CHD4 (amino acid sequence depicted on Fig.2A) using a GST tag (Fig.2B), which was cleaved off by the TEV enzyme (tobacco etch virus) before NMR binding studies were performed (purification scheme depicted on Fig.2C). Separation of the recombinant PHD2-domain from the TEV enzyme and the cleaved tag was obtained by size exclusion chromatography (SEC) and confirmed by gel electrophoresis (Fig.2D) and mass spectrometry (Fig.2E). Our mass spectrometry analysis (Fig.2E) also shows that the purified protein had the expected molecular weight (6,53 KDa) and was highly pure. The 1D ^{1}H-NMR spectrum of the PHD2-CHD4 domain (Fig.2F) shows well dispersed protein peaks which indicates that the protein was folded.

As previously stated, the PHD2 domain of CHD4 recognizes the N-terminus of histone H3 and this interaction is facilitated by methylation of Lys9[27]. Hence, to assess if the protein purified from bacteria was correctly folded and functional, we performed NMR binding assays, through protein observation, between the recombinant PHD2-CHD4 protein and the 12-mer peptides, H3K9 and H3K9me3. For this purpose, we acquired 1D ^{1}H-NMR spectra of PHD2-CHD4 in the presence of 16-fold excess of either H3K9 (Fig.3A) or H3K9me3 peptides (Fig.3B). We can observe that the presence of either of the peptides results in drastic chemical shift perturbations in the PHD2-CHD4 spectrum, which implies binding between the peptides and the recombinant protein. These results suggest that the purified PHD2 domain of CHD4 is correctly folded and functional.

^{1}H-NMR small molecule screen

After having established the purification of functional PHD2-CHD4, we performed a first small protein-compound binding screen with the 10 candidate compounds obtained from the first *in silico* screening using SEED. The screen was carried out using 30μM of recombinant protein in the presence of 16-fold excess of each compound (480μM) in protein buffer (50mM Tris-HCl, 150mM NaCl, 100μM ZnCl2, 1mM DTT, pH 8) supplemented with 10% D_2O and 0.2mM DSS (used as chemical shift reference). Before the screen was performed, 1D ^{1}H-NMR spectra were collected for each compound to determine the solubility of the compounds in protein buffer.

To assess binding, we used the same protein observation method described above to evaluate the function of the purified PHD2 domain. The 1D ^{1}H-NMR spectra acquired with the first library of 10 compounds revealed no binders (example depicted on Fig. 3C).

Previous studies[27] have shown that the binding between PHD2 and the H3 tail is weak (18 μM binding affinity towards unmodified H3 and 0.9 μM for H3K9me3). Therefore, we

decided to perform 2D $^{1}H,^{15}N$-NMR protein-compound binding experiments, as the improved signal dispersion in 2D NMR spectra is better suited to study weaker interactions.

^{15}N-labelled PHD2 is functional and shows higher affinity to methylated H3K9

To perform 2D-NMR screens we labelled the PHD2 domain with ^{15}N. The same purification procedure and quality control was performed for ^{15}N-labelled PHD2-CHD4 recombinant protein as for unlabeled protein. To evaluate the functionality of the labelled protein, we recorded 2D $^{1}H,^{15}N$-HSQC spectra of the protein and performed a titration series of both H3K9 (Fig.4A) and H3K9me3 peptides (Fig.4B) at variable peptide concentrations. Both peptides caused chemical shift perturbations in a diverse set of peaks, although the magnitude of these perturbations were significantly larger for the methylated version of H3K9. Indeed, a higher binding affinity (K_D = 0.3mM) was determined for H3K9me3, whereas a K_D of 0.7mM was measured for unmodified H3K9. These results are in agreement with previous published data[27] that showed that methylation of H3K9 increases the binding affinity between the H3 tail and the PHD2 domain. Furthermore, these data validate the use of the ^{15}N-labelled PHD2-CHD4 for compound binding assays.

^{15}N-NMR small molecule screen

A total of 105 compounds were tested for binding to PHD2-CHD4. Binding assays were performed in protein buffer (50mM Tris-HCl, 150mM NaCl, 100µM ZnCl2, 1mM DTT, pH 8, supplemented with 10%D_2O and 0.2mM DSS) with 30µM of recombinant protein and 16-fold excess of each compound (480µM). Binding was determined by chemical shift perturbations assessment by overlaying the spectrum of unbound PHD2 with the spectrum of PHD2 acquired in the presence of each compound.

As a result, no strong binders were detected, which was expected due to the loose conformation of the PHD2 domain and the weak affinity of this domain to its natural interacting partner. Weak binding was detected for compound DSP_C42_003 (Fig.5A). Chemical shift perturbations in 3 residues are illustrated in Figure 5B. The binding between this compound and the PHD2 domain of CHD4 was confirmed by ligand observation assays. One-dimensional WaterLOGSY experiments showed a positive phase for the peaks of the DSP_C42_003 compound, which indicates binding (Fig.5C). $T_1\rho$ relaxation ligand observation experiments showed an attenuation of the intensity of the compound peaks at larger relaxation delay, which is in agreement with a fast relaxation behavior, indicating binding to the protein (Fig.5D).

DSP_C42_003 possible binding interactions with PHD2-CHD4

The 4,5,6,7-tetrahydro-1H-indazole and the 1,2,4-triazole of DSP_C42_003 establish polar interactions with the carbonyl oxygen of His89 and the sidechain of Asp87. Of note, there are other potential H-bond acceptors in the vicinity, including the sidechain of Ser111 and the backbone carbonyl oxygen of Met90. Interestingly, the indazole N2 atom lies merely 0.7 Å from where the ε primary amine of H3K4 is located in the reported NMR structure (2l75.pdb).

The ethylbenzene moiety, which was not included in the docked fragment, protrudes out towards the solvent and is likely not contributing to the interaction energy. Being fully solvated, this nonpolar group can have detrimental effects on the compound's affinity by affecting its solubility or favoring spurious interactions with other proteins or molecules.

DISCUSSION

Numerous studies have demonstrated an oncogenic role for CHD4. In FP-RMS, CHD4 collaborates with the fusion protein, PAX3-FOXO1, to maintain its aberrant gene expression program[11]. In glioblastoma, this chromatin remodeler regulates expression, together with the transcription factor ZFHX4, and allows the tumor-initiation cell state[28]. In colon cancer, CHD4 keeps the expression of tumor suppressor genes silenced[29]. Furthermore, many studies have also shown that CHD4 depletion, through genetic engineering, causes tumor cell death in a variety of cancers, such as breast cancer and acute myeloid leukemia[11,13,14]. Together, these data strongly encourage the development of CHD4 inhibitors and show the vast number of possible therapeutic applications.

The enzymatic domain of CHD4, the ATPase, would be the most logical domain to use to disrupt CHD4's activity. Nevertheless, due to its complexity and size, no structure information is currently available. Moreover, all proteins from the SNF2-like family[30–32], which carry a diversity of essential physiological roles, have very similar ATPase domains, which may render development of specific inhibitors unfeasible.

NMR based structural data is available for both PHD and chromodomains of CHD4, the histone- and DNA-binding domains of the protein[33–35]. Protein-DNA binding and protein-protein interactions are difficult to model, which makes the development of inhibitors for both PHD and chromodomains challenging. Nevertheless, histone "reader" domains have been successfully targeted[5]. In fact, small molecule inhibitors of PHD domains binding to H3K4me3 have been already developed[36]. Moreover, the activity of CHD4 in transcription regulation was shown to be dependent on the binding between the PHD domains and histones[37]. Hence, we decided to interfere with CHD4's activity by targeting its second PHD domain, since PHD2 has a higher affinity than PHD1 for the H3 tail[37].

The PHD finger is a small domain characterized by a cysteine- and histidine-containing signature coordinating two Zinc ions. This domain fold is quite loose and is composed by with a short two-stranded antiparallel β-sheet, an occasional short C-terminal α-helix, and three loops of variable length and residue composition[38]. This minimal fold is maintained by the Zinc ions and conserved hydrophobic residues that form a small hydrophobic core[39]. PHD fingers bind to histone 3 N-terminal tails with a handful of features constantly observed in the available crystallographic and NMR structures, namely, (a) the backbone atoms of the first residues of the histone 3 tail establish H-bonds with the backbone of the antiparallel β-sheet of the PHD finger, effectively forming its third strand, (b) the positively charged N-terminus of the histone tail is coordinated by the oxygen atoms of the backbone carbonyl of a semi-conserved P-X-G-X-W motif (X is any residue) in the most C-terminal loop, and (c) binding does not lead to significant structural rearrangements of the conformation of the PHD finger[40,41].

The histone-binding surface of the PHD finger is shallow and polar. Therefore, while still deemed feasible targets for structure-based drug design efforts[42], PHD fingers are possibly considered the hardest epigenetic "reader" domain to drug[43-45]. Noteworthy, attempts of targeting PHD domains focused on the H3K4me3-binding subfamily of the PHD finger domains, whose members display an aromatic cage[46] which appears to be more druggable than the relatively flat, solvent exposed and polar H3K4me0-binding surface.

Given the experimental evidence and computational analyses[21–23], in our efforts to identify small molecules inhibiting the PHD2 finger/histone tail interaction, we focused on the interactions established between this domain and the first four residues of the histone tail. Our *in silico* efforts gathered a library of 105 candidate PHD2-CHD4 ligands. Validation of these candidates was performed by 2D $^{1}H,^{15}N$-NMR binding assays, based on protein observation. One compound, DSP_C42_003, showed weak binding capability towards the PHD2 domain, which was confirmed by ligand observation assays.

In spite of not being positively charged, one of the nitrogen atoms in the 4,5,6,7-tetrahydro-1H-indazole group of DSP_C42_003 seems to somehow mimic the side chain amine of H3K4, establishing polar contacts with the negatively charged patch of the PHD2 finger surface. The interactions are reinforced by similar contacts provided by another nitrogen atom of the 1,2,4-triazole. Presumably, such polar contacts in a solvent exposed area are likely to be transient and labile in nature. It will be therefore imperative for the PHD finger ligand to undergo optimization cycles in order to improve its binding affinity. To this end, experimental indications based on the protein chemical shift perturbations upon titration of the ligand and focused chemical synthesis derivatives of this hit will be crucial. For example, upon experimental validation of the calculated structure shown in Figure 6, the

first round of optimization of DSP_C42_003 could remove the ethylbenzene and branch the indazole 6-member ring introducing H-bond donors and acceptors to mimic the backbone of H3K4.

In summary, we have identified the first small molecule binder of the PHD2 domain of CHD4 and proved the feasibility of targeting of H3K9-binding "reader" domains, although further studies are required to enhance the affinity of the compound identified.

MATERIAL AND METHODS

Cloning, Expression, and Purification of PHD2-CHD4

The PHD2-CHD4 (496-551) construct was generated by PCR, using cDNA from the RH4 cell line as template and the appropriate set of primers (see supplementary material and methods). The amplified PCR product was cloned into the bacterial expression vector pETM-30, encoding a N-terminal GST-tag followed by a TEV cleavage site. Correct insertion of the PHD2-CHD4 sequence was verified by DNA sequencing. E. coli BL21 (DE3) cells were transformed with the expression vector and grown overnight at 37°C in LB media in the presence of kanamycin and 150μM of $ZnCl_2$. Then, when the culture reached an OD600 of 0.6, expression was induced with 1mM of isopropyl-thiogalactopyranoside (IPTG) for at least 4 hours. Cell pellets were frozen at -80°C and thawed in lysis buffer (50mM Tris-HCl, 150mM NaCl, 10% glycerol, 0.2% NP-40, 100μM $ZnCl_2$, 5mM DTT, pH 8, and EDTA-free protease inhibitor cocktail tablets – cOmplete Roche) on the next day. Bacteria were lysed, in the presence of DNase, by a French press. After centrifugation, the recombinant protein present in the lysate supernatant was purified by binding of the GST-tagged protein to a glutathione sepharose resin (Glutathione Sepharose 4 Fast Flow, GE Healthcare). After several washes, GST-tagged protein was eluted off the resin by 10mM of L-glutathione (Sigma Aldrich) in 50mM Tris-HCl, 150mM NaCl, 100μM $ZnCl_2$, 5mM DTT, pH 8. Subsequently, the eluted recombinant protein was submitted to a one-step overnight dialysis, where the GST-tag was cleaved off by TEV enzyme. PHD2-CHD4 was further purified by size exclusion chromatography (SEC) and concentrated through 10 KDa MWCO centrifugal concentrators (Millipore) in 50mM Tris-HCl, 150mM NaCl, 100μM $ZnCl_2$, 1mM DTT, pH 8.

^{15}N-labelled PHD2-CHD4

^{15}N-labelled PHD2-CHD4 was obtained by growing the transformed bacteria in minimal medium supplemented with $^{15}NH_4Cl$ (see supplementary material and methods). The subsequent purification was carried out following the same protocol as for the unlabeled protein.

In silico screenings

All computational screenings were performed using as target a conformation of the solution structure of PHD2-CHD4 in complex with H3K9me3 (2l75.pdb) and parametrizing the compounds' structures using CHARMM36/CGenFF force field[47,48].

In silico screenings using SEED

Rigid-ligand docking campaigns were performed using the software SEED[24,25], which calculates the poses' binding free energy using a force-field-based energy function with approximation of desolvation effects in the continuum dielectric representation. Of note, two of the NMR models (the first and the tenth) were used during this *in silico* screen to take into account the flexibility of the N-terminal tail.

First *in silico* screening using SEED

The first screen performed using SEED was run using fragments. Starting from 1,984,949 small molecules available at the time in the ZINC library, 106,534 fragments containing at least a ring and a H-bond donor or acceptor were generated using DAIM[49], and 96,980 of which were parametrized using MATCH[50]. The overall procedure can be found in[51].

These fragments were docked into the PHD2-CHD4 domain. SEED allows the user to provide a list of coordinates that serve to select polar and apolar receptor vectors. For this docking, three coordinates of the histone tail residues H3T3 and H3K4 lying in the center of the histone tail-binding region were used. The docked fragments were then scored according to a consensus calculated as the median of the rankings of the following terms: receptor desolvation, van der Waals, and ligand efficiency.

The 200-best scoring fragments were visually inspected and, in case the parent compound had no moieties sterically interfering with the calculated binding mode, the compound was selected. This resulted in the selection and purchase of 40 and 10 compounds, respectively.

Second *in silico* screening using SEED

514 compounds, available at the time at the Biomolecular Screening Facility of the EPFL in Lausanne with no more than two rotatable bonds, were selected as starting point for the *in silico* library of the second computational screening using SEED. 3,464 tautomers were generated using the calculator plugin of Marvin and docked using SEED.

Docking and scoring approaches similar to those described previously were employed which ultimately resulted in the purchase of 72 compounds for experimental validation.

In silico screening using Pharmit

In parallel, we undertook a pharmacophore search using Pharmit[26]. Based on the key mediators of the PHD finger/histone peptide interactions we chose four pharmacophores, two polar (the H-bond donor and acceptor based on H3R2 carbonyl oxygen and the H3K4 backbone amide hydrogen) and two non-polar (the side chain methyl groups of H3A1 and H3T3). Screens were run using seven libraries (namely CHEMBL21, ChemDiv, Chemspace, MolPort, NCI Open Chemical Repository, PubChem, and ZINC), resulting in more than 145,000,000 compounds. The exclusive shape constraints was set to 0.5, to implicitly take into account the molecules' flexibility, and an energy minimization step was performed to optimize shape complementarity and H-bonds. The filtering based on the molecules' properties (non-hydrogen atom counts lower than 26 and logP lower than 3) resulted in 9,340 compounds.

Then, 24,730 tautomers were generated taking advantage of the calculator plugin of Marvin, keeping only the 9,497 that were represented in at least 40% of the conformations at pH 7.4. The 4,624 molecules containing at least one positive charge were then parametrized and the PHD finger/small molecule complex was energy-minimized. The *pbeq* plugin of CHARMM[52] was used for finite-difference Poisson calculations on the minimized structures[53].

To analyze the plethora of calculated binding poses, the following filters were applied: (1) the computed strain energy for the small molecule had to be lower than 5 kcal/mol, (2) a nitrogen atom had to be present within 8.5 Å of the point where the N-terminus of the H3 peptide is supposed to be, and (3) compounds had not to contain nitro groups. 3,404 compounds were left after enforcing these filters. Several consensus scores were then calculated based on different permutations of a subset of the following criteria: (a) CHARMM total energy, (b) intermolecular electrostatic term normalized on the heavy atom count, (c) intermolecular van der Waals term normalized on the heavy atom count, (d) CHARMM total energy normalized on the heavy atom number count, and (e) difference between the electrostatic interaction energy between ligand and protein in water and the free energy of solvation of the ligand in the conformation upon investigation. Upon visual inspection of the best-scoring conformations, 23 compounds were selected.

NMR spectroscopy

All NMR spectra were acquired at 296K on 600 or 800MHz Bruker spectrometers (Bruker, Switzerland) equipped with a 5mm TXI triple resonance Cryoprobe™ and a SampleJet™ for sample automation, which was fed with 3mm NMR tubes containing a sample volume of

180µl. The NMR spectra were transformed and analyzed with TopSpin (Bruker, Switzerland).

Solubility assays

The solubility of all tested compounds was assessed by 1D ^{1}H-NMR spectra, which were also used to confirm the structure of the purchased compounds. All compounds were first dissolved in d_6-DMSO to a 50mM concentration. Then, they were diluted in protein buffer (50mM Tris-HCl, 150mM NaCl, 100µM $ZnCl_2$, 1mM DTT, pH 8) to a final concentration of 0.6mM in the presence of 10% $^{2}H_2O$ and 0.2mM DSS (4,4-dimethyl-4-silapentane-1-sulfonic acid), used as an internal concentration and chemical shift reference. The actual compound concentration was determined from the integral ratio of the compound peaks and the 0.00 ppm DSS signal.

Ligand-protein binding experiments

Ligand-protein binding experiments were carried out either with unlabeled or uniformly ^{15}N-labelled PHD2-CHD4 in protein buffer (as described above) supplemented with 10% $^{2}H_2O$ and 0.2mM DSS. For all experiments a protein concentration of 30µM was used. All compounds were individually tested in a 16-fold excess relative to the protein concentration (480µM). Spectra of the protein in the presence of an equivalent amount of d_6-DMSO were also acquired as a control. The 12-mer peptides H3K9 and H3K9me3, the natural PHD2-CHD4 interactors, were purchased from LifeTein LLC, Hillsborough, NJ, USA. They were used as positive controls and to assess the functionality of the produced recombinant PHD2-CHD4 protein.

For NMR screening, series of ligand observation 1D ^{1}H $T_1\rho$ relaxation experiments[54] with spin-lock times of 10ms and 200ms, and PO-WaterLOGSY experiments[55] were performed. In addition to the experiments in the presence of the protein, as described above, the same measurements were performed in the absence of protein as negative controls. A compound was identified as a hit in the case that one or more of its lines in the spectra recorded in the presence of protein showed a signal decrease of at least 33% in the $T_1\rho$ relaxation spectrum with 200ms spin-lock duration, when compared with the one recorded with 10ms ($I_{200} : I_{10} < 0.67$). Additionally, these changes in signal intensity cannot not occur in the corresponding spectra recorded in absence of protein (negative control). Moreover, hits were identified by observation of positive signals in WaterLOGSY spectra (relative to Tris and DMSO signals, which were arbitrarily phased with negative sign).

In addition to the ligand observation experiments described above, all compounds were screened as singles by protein observation NMR experiments. These experiments were used

to assess chemical shift perturbations by compound binding, using either unlabeled protein to measure 1D ^{1}H-NMR spectra or ^{15}N-labelled PHD2-CHD4 to record 2D ^{15}N-HSQC (heteronuclear single quantum coherence) spectra.

Binding affinities with the natural interactor peptides, H3K9 and H3K9me3, were determined by monitoring the chemical shift changes of the backbone amide as a function of peptide concentration. Multiple binding curves were obtained by titrating each peptide into separate samples of ^{15}N-labelled PHD2-CHD4 and acquiring 2D ^{15}N-HSQC spectra for 4 different ligand:protein ratios from the range 1:1 to 16:1. Chemical shift changes in both ^{1}H and ^{15}N dimensions were logged and fitted by non-linear regression analysis to a single site binding model, using a least square fitting search to find the values of K_D and the chemical shift of the fully saturated protein[56].

ACNOWLEDGMENTS

We would like to thank Chrystele Henry and Suzanne Chau for their help in the acquirement of NMR spectra. Also, we would like to thank the Novartis Institutes for Biomedical Research for allowing us to use their spectrometers. Finally, we would like to thank Prof. Rudi Glockshuber for granting us access to the protein purification facility present in his laboratory.

REFERENCES

1. Annunziato, A. DNA Packaging: Nucleosomes and Chromatin | Learn Science at Scitable. *Nature Education* (2008). doi:1(1):26
2. Luger, K., Mäder, A. W., Richmond, R. K., Sargent, D. F. & Richmond, T. J. Crystal structure of the nucleosome core particle at 2.8 Å resolution. *Nature* **389,** 251–260 (1997).
3. Lawrence, M., Daujat, S. & Schneider, R. Lateral Thinking: How Histone Modifications Regulate Gene Expression. (2016). doi:10.1016/j.tig.2015.10.007
4. Kouzarides, T. Chromatin Modifications and Their Function. *Cell* **128,** 693–705 (2007).
5. Filippakopoulos, P. *et al.* Selective inhibition of BET bromodomains. *Nature* **468,** 1067–1073 (2010).
6. Portela, A. & Esteller, M. Epigenetic modifications and human disease. *Nat. Biotechnol.* **28,** 1057–1068 (2010).
7. Erez-Salvia, M. P. & Esteller, M. Bromodomain inhibitors and cancer therapy: From structures to applications. *Epigenetics* **12,** (2017).
8. Musselman, C. A. & Kutateladze, T. G. PHD Fingers: Epigenetic Effectors and Potential

Drug Targets. *Mol. Interv.* **9,** 314–323 (2009).

9. Hall, J. A. & Georgel, P. T. CHD proteins: a diverse family with strong ties. *Biochem. Cell Biol.* **85,** 463–476 (2007).
10. Xue, Y. *et al.* NURD, a Novel Complex with Both ATP-Dependent Chromatin-Remodeling and Histone Deacetylase Activities. *Mol. Cell* **2,** 851–861 (1998).
11. Böhm, M. *et al.* Helicase CHD4 is an epigenetic coregulator of PAX3-FOXO1 in alveolar rhabdomyosarcoma. *J. Clin. Invest.* **126,** 4237–4249 (2016).
12. Sankar, S. *et al.* Mechanism and relevance of EWS/FLI-mediated transcriptional repression in Ewing sarcoma. *Oncogene* **32,** 5089–100 (2013).
13. Heshmati, Y. *et al.* Identification of CHD4 As a Potential Therapeutic Target of Acute Myeloid Leukemia. *Blood* **128,** (2016).
14. D'Alesio, C. *et al.* RNAi screens identify CHD4 as an essential gene in breast cancer growth. *Oncotarget* **7,** 80901–80915 (2016).
15. Bouazoune, K. *et al.* The dMi-2 chromodomains are DNA binding modules important for ATP-dependent nucleosome mobilization. *EMBO J.* **21,** 2430–40 (2002).
16. Tencer, A. H. *et al.* Covalent Modifications of Histone H3K9 Promote Binding of CHD3. *Cell Rep.* **21,** 455–466 (2017).
17. Mansfield, R. E. *et al.* Plant Homeodomain (PHD) Fingers of CHD4 Are Histone H3-binding Modules with Preference for Unmodified H3K4 and Methylated H3K9. *J. Biol. Chem.* **286,** 11779–11791 (2011).
18. Musselman, C. A. *et al.* Binding of the CHD4 PHD2 finger to histone H3 is modulated by covalent modifications. *Biochem. J.* **423,** 179–187 (2009).
19. Watson, A. A. *et al.* The PHD and Chromo Domains Regulate the ATPase Activity of the Human Chromatin Remodeler CHD4. *J. Mol. Biol.* **422,** 3–17 (2012).
20. Morra, R., Lee, B. M., Shaw, H., Tuma, R. & Mancini, E. J. Concerted action of the PHD, chromo and motor domains regulates the human chromatin remodelling ATPase CHD4. *FEBS Lett.* **586,** 2513–2521 (2012).
21. Chignola, F. *et al.* The solution structure of the first PHD finger of autoimmune regulator in complex with non-modified histone H3 tail reveals the antagonistic role of H3R2 methylation. *Nucleic Acids Res.* **37,** 2951–61 (2009).
22. Org, T. *et al.* The autoimmune regulator PHD finger binds to non-methylated histone H3K4 to activate gene expression. *EMBO Rep.* **9,** 370–6 (2008).
23. Spiliotopoulos, D., Spitaleri, A. & Musco, G. Exploring PHD fingers and H3K4me0 interactions with molecular dynamics simulations and binding free energy calculations: AIRE-PHD1, a comparative study. *PLoS One* **7,** e46902 (2012).
24. Majeux, N., Scarsi, M. & Caflisch, A. Efficient electrostatic solvation model for protein-

fragment docking. *Proteins* **42,** 256–68 (2001).

25. Majeux, N., Scarsi, M., Apostolakis, J., Ehrhardt, C. & Caflisch, A. Exhaustive docking of molecular fragments with electrostatic solvation. *Proteins* **37,** 88–105 (1999).
26. Sunseri, J. & Koes, D. R. Pharmit: interactive exploration of chemical space. *Nucleic Acids Res.* **44,** W442-8 (2016).
27. Musselman, C. A. *et al.* Binding of the CHD4 PHD2 finger to histone H3 is modulated by covalent modifications. *Biochem. J.* **423,** 179–187 (2009).
28. Chudnovsky, Y. *et al.* ZFHX4 Interacts with the NuRD Core Member CHD4 and Regulates the Glioblastoma Tumor-Initiating Cell State. *Cell Rep.* **6,** 313–324 (2014).
29. Xia, L. *et al.* CHD4 Has Oncogenic Functions in Initiating and Maintaining Epigenetic Suppression of Multiple Tumor Suppressor Genes. *Cancer Cell* **31,** 653–668.e7 (2017).
30. Flaus, A., Martin, D. M. A., Barton, G. J. & Owen-Hughes, T. Identification of multiple distinct SNF2 subfamilies with conserved structural motifs. *Nucleic Acids Res.* **34,** 2887–2905 (2006).
31. Dürr, H., Flaus, A., Owen-Hughes, T. & Hopfner, K.-P. SNF2 family ATPases and DExx box helicases: differences and unifying concepts from high-resolution crystal structures. *Nucleic Acids Res.* **34,** 4160–7 (2006).
32. Eisen, J. A., Sweder, K. S. & Hanawalt, P. C. Evolution of the SNF2 family of proteins: subfamilies with distinct sequences and functions. *Nucleic Acids Res.* **23,** 2715–23 (1995).
33. Mansfield, R. E. *et al.* Plant homeodomain (PHD) fingers of CHD4 are histone H3-binding modules with preference for unmodified H3K4 and methylated H3K9. *J. Biol. Chem.* **286,** 11779–91 (2011).
34. Wai, D. C. C. *et al.* The BRD3 ET domain recognizes a short peptide motif through a mechanism that is conserved across chromatin remodelers and transcriptional regulators. *J. Biol. Chem.* **293,** 7160–7175 (2018).
35. RCSB PDB - Search Results. Available at: http://www.rcsb.org/pdb/results/results.do?tabtoshow=Current&qrid=18DBE8A4. (Accessed: 13th July 2018)
36. Wagner, E. K., Nath, N., Flemming, R., Feltenberger, J. B. & Denu, J. M. Identification and Characterization of Small Molecule Inhibitors of a Plant Homeodomain Finger. *Biochemistry* **51,** 8293–8306 (2012).
37. Musselman, C. A. *et al.* Bivalent recognition of nucleosomes by the tandem PHD fingers of the CHD4 ATPase is required for CHD4-mediated repression. *Proc. Natl. Acad. Sci. U. S. A.* **109,** 787–92 (2012).

38. Baker, L. A., Allis, C. D. & Wang, G. G. PHD fingers in human diseases: disorders arising from misinterpreting epigenetic marks. *Mutat. Res.* **647,** 3–12 (2008).
39. Musco, G. & Peterson, P. PHD finger of autoimmune regulator: an epigenetic link between the histone modifications and tissue-specific antigen expression in thymus. *Epigenetics* **3,** 310–4 (2008).
40. Sanchez, R. & Zhou, M.-M. The PHD finger: a versatile epigenome reader. *Trends Biochem. Sci.* **36,** 364–72 (2011).
41. Musselman, C. A. & Kutateladze, T. G. Handpicking epigenetic marks with PHD fingers. *Nucleic Acids Res.* **39,** 9061–71 (2011).
42. Arrowsmith, C. H., Bountra, C., Fish, P. V, Lee, K. & Schapira, M. Epigenetic protein families: a new frontier for drug discovery. *Nat. Rev. Drug Discov.* **11,** 384–400 (2012).
43. Santiago, C., Nguyen, K. & Schapira, M. Druggability of methyl-lysine binding sites. *J. Comput. Aided. Mol. Des.* **25,** 1171–8 (2011).
44. Miller, T. C. R. *et al.* Competitive binding of a benzimidazole to the histone-binding pocket of the Pygo PHD finger. *ACS Chem. Biol.* **9,** 2864–74 (2014).
45. Wagner, E. K., Nath, N., Flemming, R., Feltenberger, J. B. & Denu, J. M. Identification and characterization of small molecule inhibitors of a plant homeodomain finger. *Biochemistry* **51,** 8293–306 (2012).
46. Taverna, S. D., Li, H., Ruthenburg, A. J., Allis, C. D. & Patel, D. J. How chromatin-binding modules interpret histone modifications: lessons from professional pocket pickers. *Nat. Struct. Mol. Biol.* **14,** 1025–1040 (2007).
47. MacKerell, A. D., Feig, M. & Brooks, C. L. Improved treatment of the protein backbone in empirical force fields. *J. Am. Chem. Soc.* **126,** 698–9 (2004).
48. Vanommeslaeghe, K. *et al.* CHARMM general force field: A force field for drug-like molecules compatible with the CHARMM all-atom additive biological force fields. *J. Comput. Chem.* **31,** 671–90 (2010).
49. Kolb, P. & Caflisch, A. Automatic and efficient decomposition of two-dimensional structures of small molecules for fragment-based high-throughput docking. *J. Med. Chem.* **49,** 7384–92 (2006).
50. Yesselman, J. D., Price, D. J., Knight, J. L. & Brooks, C. L. MATCH: an atom-typing toolset for molecular mechanics force fields. *J. Comput. Chem.* **33,** 189–202 (2012).
51. Xu, M. *et al.* Discovery of CREBBP Bromodomain Inhibitors by High-Throughput Docking and Hit Optimization Guided by Molecular Dynamics. *J. Med. Chem.* **59,** 1340–9 (2016).
52. Brooks, B. R. *et al.* CHARMM: the biomolecular simulation program. *J. Comput. Chem.*

30, 1545–614 (2009).

53. Marchand, J.-R., Dalle Vedove, A., Lolli, G. & Caflisch, A. Discovery of Inhibitors of Four Bromodomains by Fragment-Anchored Ligand Docking. *J. Chem. Inf. Model.* **57,** 2584–2597 (2017).
54. Philip J. Hajduk, Edward T. Olejniczak, and & Fesik*, S. W. One-Dimensional Relaxation- and Diffusion-Edited NMR Methods for Screening Compounds That Bind to Macromolecules. (1997). doi:10.1021/JA9715962
55. Gossert, A. D., Henry, C., Blommers, M. J. J., Jahnke, W. & Fernández, C. Time efficient detection of protein–ligand interactions with the polarization optimized PO-WaterLOGSY NMR experiment. *J. Biomol. NMR* **43,** 211–217 (2009).
56. Fielding, L. NMR methods for the determination of protein-ligand dissociation constants. *Curr. Top. Med. Chem.* **3,** 39–53 (2003).

FIGURES

Figure 1

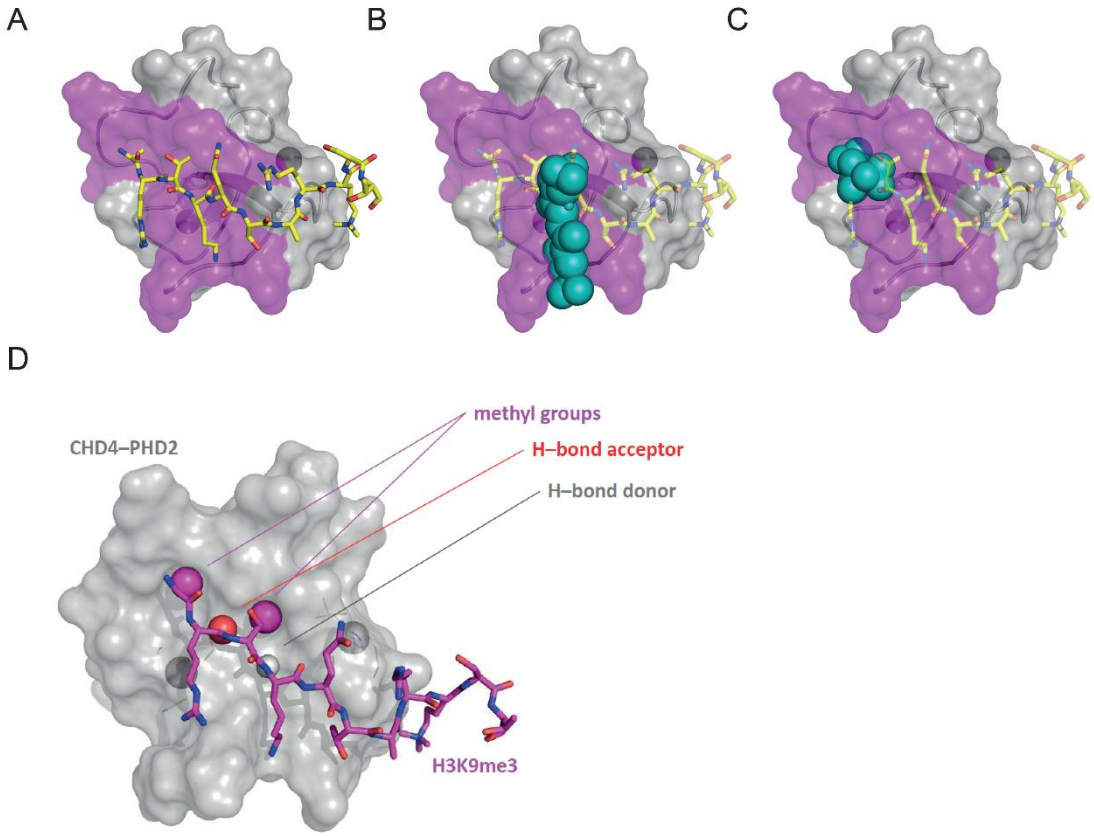

Fig.1 – Molecular model of the second PHD domain of CHD4 binding to H3K9me3 allows *in silico* screening of potential small molecule binders
(A-C) Molecular model of PHD2 domain of CHD4 (PDB code: 2L75), with the PHD finger shown as a grey cartoon and a grey surface (the grey spheres indicate the Zinc atoms) and the H3K9me3 peptide indicated by sticks (C, N and O atoms are colored in yellow, blue, and red, respectively). The areas targeted by the *in silico* screenings performed using SEED are colored in purple. Two possible CHD4-PHD2 small molecule binders, DSP_C42_004 (B) and DSP_C42_008 (C), are shown as teal spheres. (D) The pharmacophores chosen for the in silico screen performed using Pharmit are shown with spheres (violet, red, and white for "methyl groups", "H-bond acceptors" and "H-bond donors", respectively). The PHD finger is shown as grey surface (grey spheres represent the Zinc atoms) while the H3K9me3 peptide is shown as sticks (C, N and O atoms colored in violet, blue and red, respectively).

Figure 2

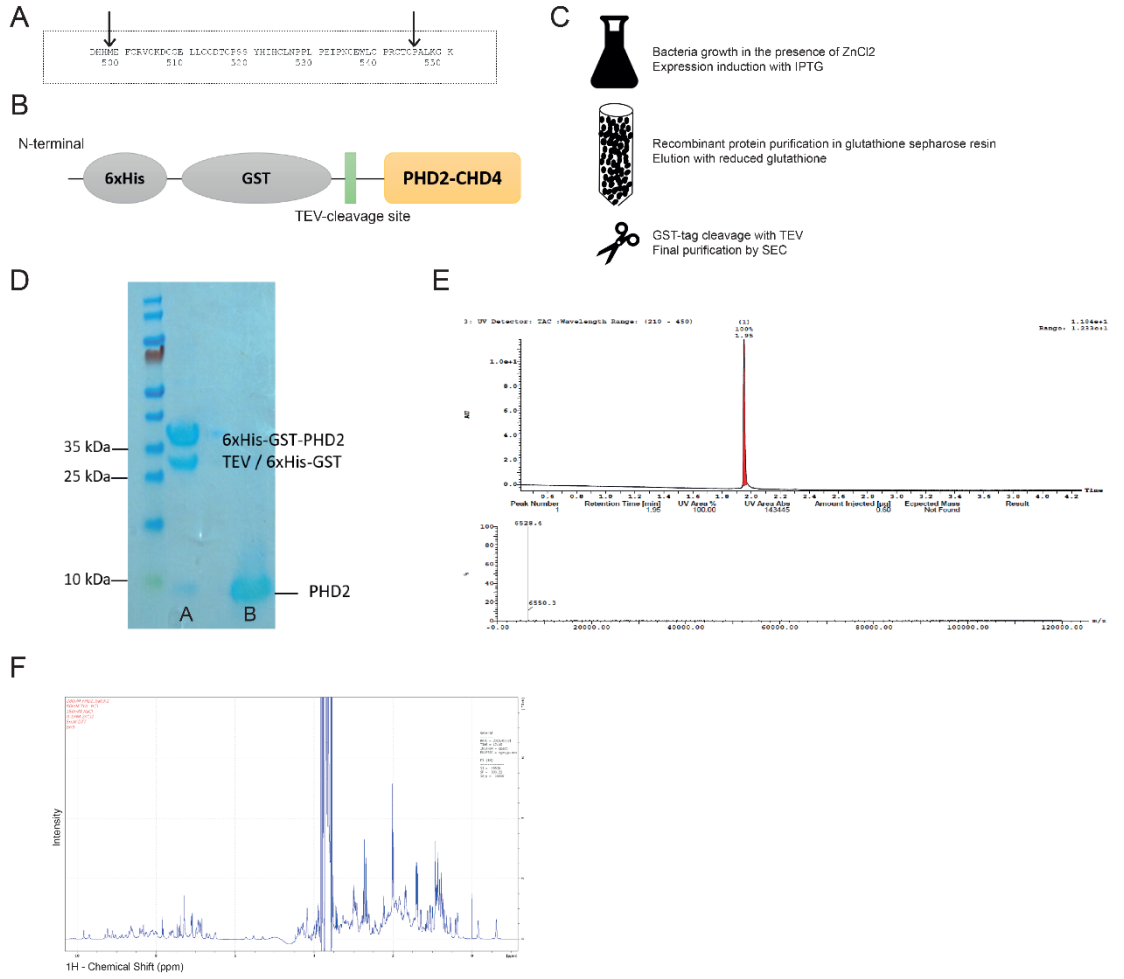

Fig.2 – Purification of the PHD2 domain of CHD4
(A) Amino acid sequence of CHD4 (496-551) cloned into the expression construct. Arrows indicate the beginning and end of the PHD2 domain. (B) Scheme of the recombinant GST-tagged PHD2-CHD4 protein. (C) Purification scheme of the PHD2 domain of CHD4. (D) Gel electrophoresis demonstrates the purity of the obtained recombinant PHD2-CHD4. In A, we can observe the bands corresponding to the GST-tagged PHD2, the TEV enzyme, the cleaved tag, and the recombinant PHD2-CHD4. B shows the purity of the recombinant protein obtained after SEC. (E) Mass spectrometry confirms the purity of the recombinant PHD2-CHD4. (F) One-dimensional ^{1}H-NMR spectrum of purified PHD2-CHD4 (30µM) recorded in 50mM Tris-HCl, 150mM NaCl, 100µM $ZnCl_2$, 1mM DTT, pH 8, supplemented with 10% $^{2}H_2O$ and 0.2mM DSS.

Figure 3

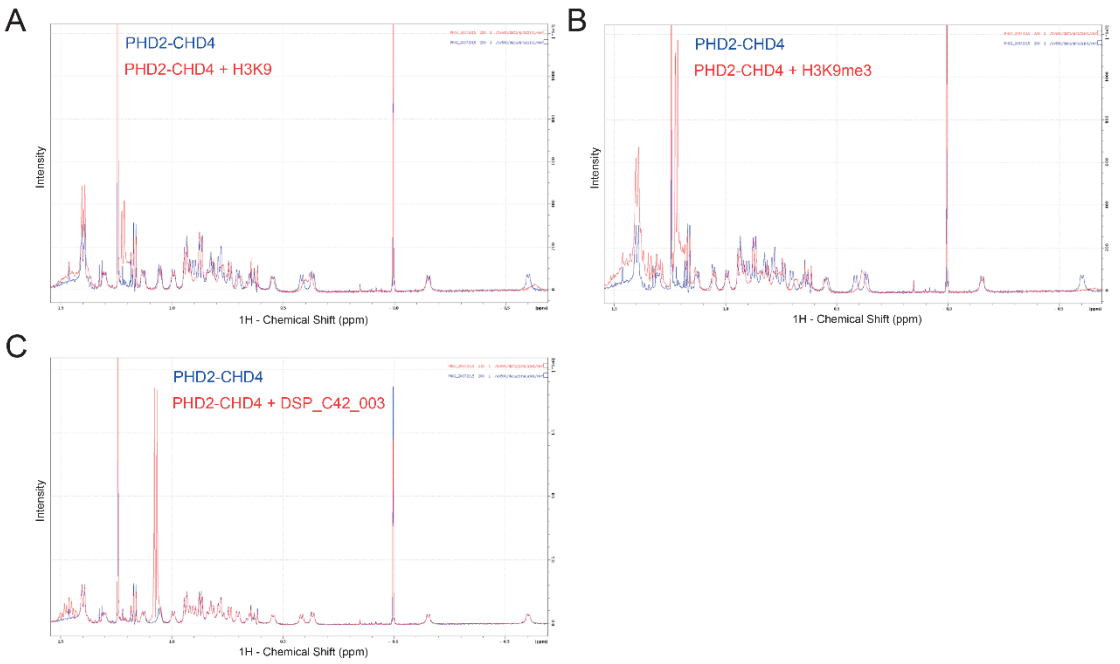

Fig.3 – Purified PHD2-CHD4 can bind 12-mer H3K9 and H3K9me3 peptides
(A) One-dimensional ^{1}H-NMR spectrum of purified PHD2-CHD4 (30µM) in the presence (red) or absence (blue) of the 12-mer H3K9 peptide (480µM). (B) One-dimensional ^{1}H-NMR spectrum of purified PHD2-CHD4 (30µM) in the presence (red) or absence (blue) of the 12-mer H3K9me3 peptide (480µM). (C) Example of first compound screen. One-dimensional ^{1}H-NMR spectrum of purified PHD2-CHD4 (30µM) in the presence (red) or absence (blue) of 480µM of DSP_C42_003.

Figure 4

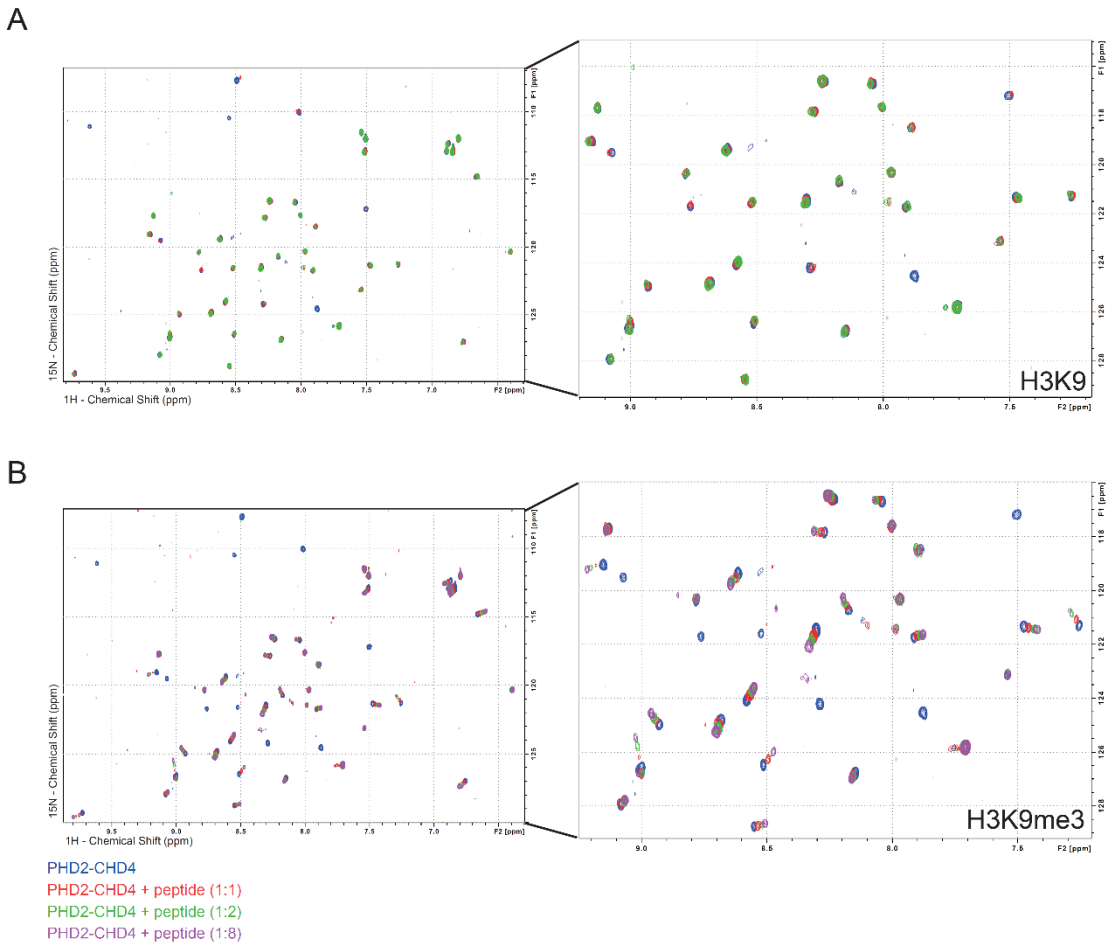

Fig.4 – Purified ^{15}N-labelled PHD2-CHD4 binds to H3K9 with increased affinity when Lys9 is methylated
(A and B) Overlay of the 2D ^{1}H,^{15}N-HSQC spectra of unbound PHD2 (blue) and PHD2 in the presence of increasing concentrations (30μM – red, 60μM – green, 120μM – purple) of the 12-mer H3K9 (A) and H3K9me3 (B) peptides.

Figure 5

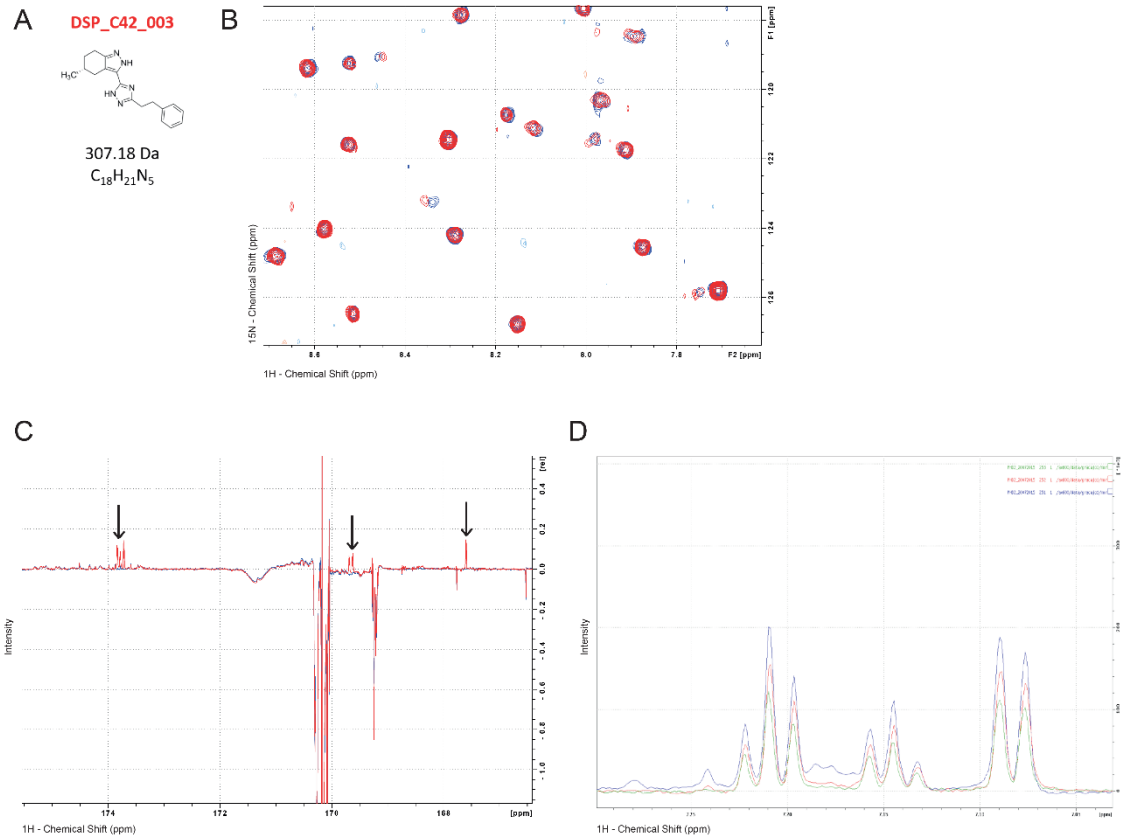

Fig.5 – DSP_C42_003 binds weakly to the PHD2 domain of CHD4
(A) Chemical structure of DSP_C42_003. (B) Overlay of the 2D ^{1}H,^{15}N-HSQC spectra of 30µM of unbound PHD2 (blue) and of 30µM of PHD2 in the presence of 480µM of DSP_C42_003 (red). (C) One-dimensional WaterLOGSY spectra of 30µM of PHD2 (in blue) and of 30µM of PHD2 in the presence of 480µM of DSP_C42_003. The arrows in the spectra indicate the positive peaks of the compound. (D) $T_1\rho$ shows binding between the PHD2 domain of CHD4 and compound DSP_C42_003. The one-dimensional ^{1}H-NMR spectrum of 480µM DSP_C42_003 in the presence of 30µM of PHD2 is shown in blue. In green (200ms) and red (10ms) are shown the one-dimensional ^{1}H-NMR spectra of 480µM DSP_C42_003 in the presence of 30µM of PHD2.

Figure 6

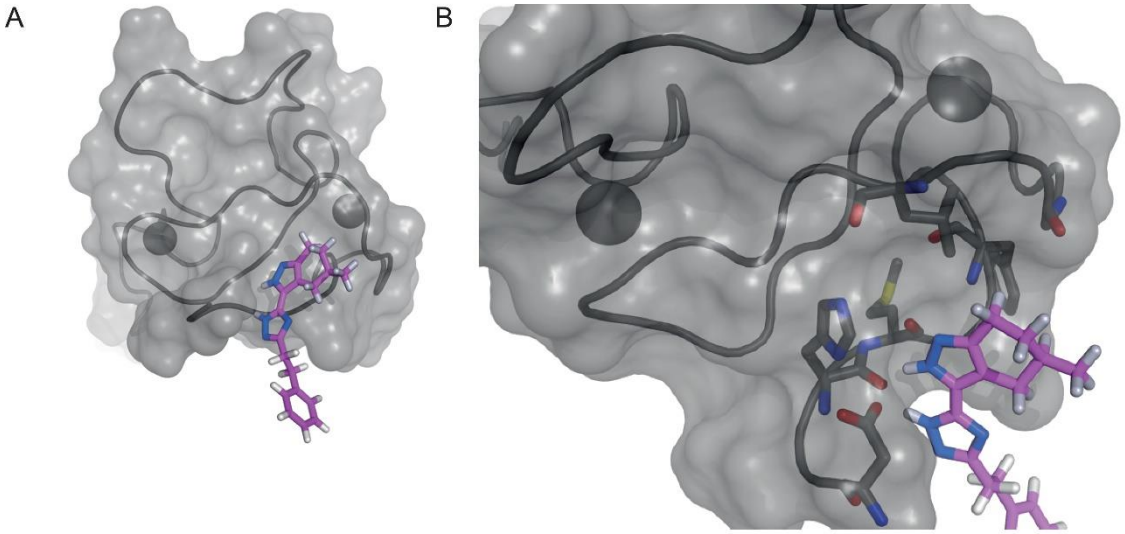

Fig.6 – Molecular modeling of DSP_C42_003 reveals residues potentially important for its binding to PHD2
(A) Docked binding pose of DSP_C42_003 in the second PHD finger of CHD4. The protein domain is shown as in Figure 1A, while the small molecule is shown as sticks (C atoms are colored in violet, N in blue, and H in grey). (B) Zoom of the docked binding pose, with relevant residues of the PHD finger shown as sticks

SUPPLEMENTARY MATERIAL AND METHODS

Cloning primers

ETM_PHD2_FW – TTTCAGGGCGCCATGGACCACCATATGGAATTCTG

ETM_PHD2_RV - GACGGAGCTCGAATTCTATTTGCCCTTCAGAGCTG

Minimal media composition

Supplementary Table 1 Minimal media components

Minimal Media	Concentration
trace elements	1x
NH_4Cl	1g/L
Glucose	0.4%
$MgSO_4$	0.2µM
$CaCl_2$	0.3µM
Biotin	1mg/L
Thiamin	1mg/L
$ZnCl_2$	0.15mM
5X M9	1x

Supplementary Table 2 100x trace elements composition

100x trace elements	Concentration (mM)
EDTA	17.1
$FeCl_3*6H_2O$	3.07
$ZnCl_2$	0.62
$CuCl_2*2H_2O$	0.08
$CoCl_2*6H_2O$	0.04
H_3BO_3	0.16
$MnCl_2*4H_2O$	0.007

Supplementary Table 3 M9 media composition

5x M9	Concentration (M)
Na_2HPO_4	0.250
KH_2PO_4	0.110
NaCl	0.04

SUPPLEMENTARY FIGURES

Supplementary Figure 1

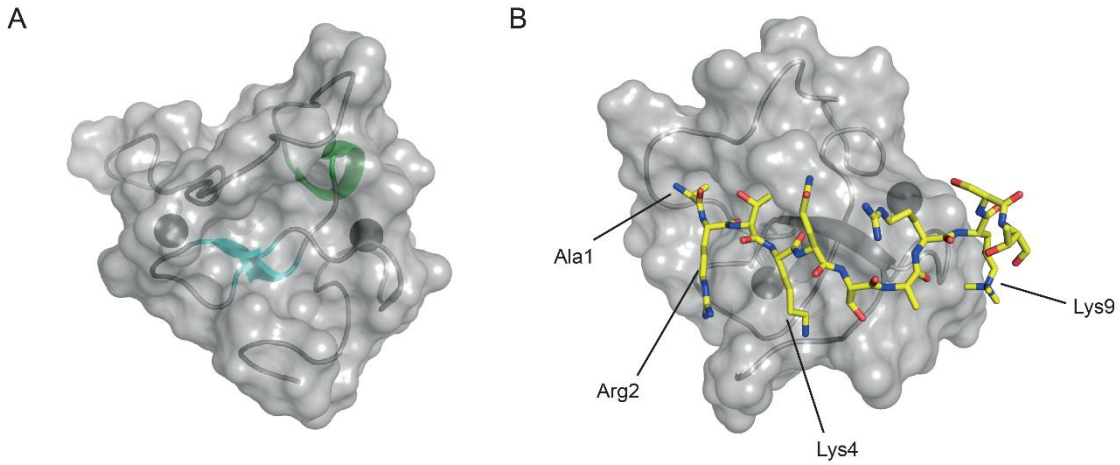

Fig.S1 – Molecular model of the second PHD domain of CHD4 and its interaction with the first nine amino acids of the H3 N-tail
(A) Molecular model of PHD2 domain of CHD4 (PDB code: 2L75), with the PHD finger shown as a grey cartoon and a grey surface. In green we can observe the small α-helix and in blue the β-sheet. Also visible are the two zinc ions which are coordinated with this domain in a cross-brace arrangement. (B) Molecular model of the interaction between the PHD2 domain and the first nine amino acids of the H3 N-tail, shown as sticks (C atoms are in yellow, N in blue and O in red). Indicated are the main amino acids responsible for the binding between PHD2 and H3.

Manuscript 3

Helicase CHD4 is an epigenetic coregulator of PAX3-FOXO1 in alveolar rhabdomyosarcoma

Maria Böhm[1], Marco Wachtel[1], Joana G. Marques[1], Natalie Streiff[1], Dominik Laubscher[1], Paolo Nanni[2], Kamel Mamchaoui[3], Raffaella Santoro[4] and Beat W. Schäfer[1]

[1]Department of Oncology and Children's Research Center, University Children's Hospital Zurich, Zurich, Switzerland
[2]Functional Genomics Center Zurich, University of Zurich, Zurich, Switzerland
[3]Thérapie des maladies du muscle strié, Institut de Myologie, UM76, UPMC Université Paris, France
[4]Institute of Veterinary Biochemistry and Molecular Biology, University of Zurich, Zurich, Switzerland

Conflict of Interest: The authors declare no conflict of interest.

Published manuscript.

Corresponding Author: Beat W. Schaefer

Department of Oncology, Children's Hospital Zurich
Steinwiesstrasse 75
8032 Zurich, Switzerland
Beat.schaefer@kispi.uzh.ch
Phone: +41 (44) 266 7553
Fax: +41 (44) 634 8859

Personal contribution: I performed all the ChIP-qPCR experiments (Fig.6B/C), analyzed the results and interpreted data. Also, I reviewed the paper and made contributions to its intellectual content.

Helicase CHD4 is an epigenetic coregulator of PAX3-FOXO1 in alveolar rhabdomyosarcoma

Maria Böhm,[1] Marco Wachtel,[1] Joana G. Marques,[1] Natalie Streiff,[1] Dominik Laubscher,[1] Paolo Nanni,[2] Kamel Mamchaoui,[3] Raffaella Santoro,[4] and Beat W. Schäfer[1]

[1]Department of Oncology and Children's Research Center, University Children's Hospital Zurich, Zurich, Switzerland. [2]Functional Genomics Center Zurich, University of Zurich, Zurich, Switzerland. [3]Sorbonne Universités, Université Pierre et Marie Curie (UPMC), Université de Paris, INSERM UMRS 974, CNRS FRE3617, Center for Research in Myology, Paris, France. [4]Institute of Veterinary Biochemistry and Molecular Biology, University of Zurich, Zurich, Switzerland.

A vast number of cancer genes are transcription factors that drive tumorigenesis as oncogenic fusion proteins. Although the direct targeting of transcription factors remains challenging, therapies aimed at oncogenic fusion proteins are attractive as potential treatments for cancer. There is particular interest in targeting the oncogenic PAX3-FOXO1 fusion transcription factor, which induces alveolar rhabdomyosarcoma (aRMS), an aggressive cancer of skeletal muscle cells for which patient outcomes remain dismal. In this work, we have defined the interactome of PAX3-FOXO1 and screened 60 candidate interactors using siRNA-mediated depletion to identify candidates that affect fusion protein activity in aRMS cells. We report that chromodomain helicase DNA binding protein 4 (CHD4), an ATP-dependent chromatin remodeler, acts as crucial coregulator of PAX3-FOXO1 activity. CHD4 interacts with PAX3-FOXO1 via short DNA fragments. Together, they bind to regulatory regions of PAX3-FOXO1 target genes. Gene expression analysis suggested that CHD4 coregulatory activity is essential for a subset of PAX3-FOXO1 target genes. Depletion of CHD4 reduced cell viability of fusion-positive but not of fusion-negative RMS in vitro, which resembled loss of PAX3-FOXO1. It also caused specific regression of fusion-positive xenograft tumors in vivo. Therefore, this work identifies CHD4 as an epigenetic coregulator of PAX3-FOXO1 activity, providing rational evidence for CHD4 as a potential therapeutic target in aRMS.

Introduction

Rhabdomyosarcoma (RMS) is a heterogeneous family of tumors sharing features with the skeletal muscle lineage. It is the most common soft tissue sarcoma in childhood and can be classified into 2 main histological subtypes, alveolar RMS (aRMS) and embryonal RMS (eRMS). These are not only histologically distinguishable but also associated with distinct clinical and molecular profiles. aRMS is the more aggressive form of RMS, with a higher rate of metastasis and a poorer prognosis. Current treatment modalities, including chemotherapy, surgery, and radiation, have steadily improved survival of RMS patients, but, especially for aRMS patients with metastatic disease, survival rates remain dismal (1, 2). So far there are no targeted therapies available that could improve overall cure rates and reduce long-term sequelae. About 80% of aRMSs are characterized by a specific chromosomal translocation generating PAX3-FOXO1 or PAX7-FOXO1 fusion transcription factors (fusion-positive RMS [FP-RMS]) (3). It is now well accepted that fusion status drives unfavorable outcome in patients with RMS, especially for the PAX3-FOXO1 fusion (4–7). Therefore fusion status has become a very important prognostic marker in the clinics. The fusion proteins are characterized by aberrant expression levels, greater posttranslational stability, exclusive nuclear localization, and a more potent transcriptional function compared with WT PAX3/PAX7 (8–11). Furthermore, numerous studies have highlighted the oncogenic ability of the PAX3-FOXO1 fusion protein and its fundamental contribution toward FP-RMS tumorigenesis (12–17). Since FP-RMS cells are addicted to the oncogenic capacity of PAX3-FOXO1 expression, it remains a very attractive therapeutic target (18–21). However, direct targeting of transcription factors persists in being challenging. Therefore, many studies have been conducted to identify PAX3-FOXO1 downstream targets or cooperating mutations that are potentially required for oncogenic transformation (reviewed in ref. 22). Many putative therapeutically relevant signaling pathways have been described with varying degrees of tumor cells' dependence on them and therefore also varying tumor response upon inhibition. Also, their complexity, crosstalk, and acquired drug resistance often limit clinical application of drugs targeting key components of these pathways (23–25). Furthermore, pediatric cancers in general are known to carry very few mutations compared with adult tumors, and recent comprehensive genomic analysis has identified a particularly low overall mutational burden in FP-RMS, decreasing the number of potential actionable targets and underscoring the importance of the fusion proteins as dominant driver (26, 27).

Epigenetic regulation is an emerging field in cancer biology and provides novel therapeutic possibilities (28). Several histone demethylases and histone methyltransferases are highly expressed in FP-RMS and have been shown to maintain the undifferentiated phenotype of tumor cells or promote their survival. In these known cases, epigenetic modulators act as either direct

Conflict of interest: The authors have declared that no conflict of interest exists.
Submitted: October 13, 2015; **Accepted:** September 8, 2016.
Reference information: *J Clin Invest*. 2016;126(11):4237–4249. doi:10.1172/JCI85057.

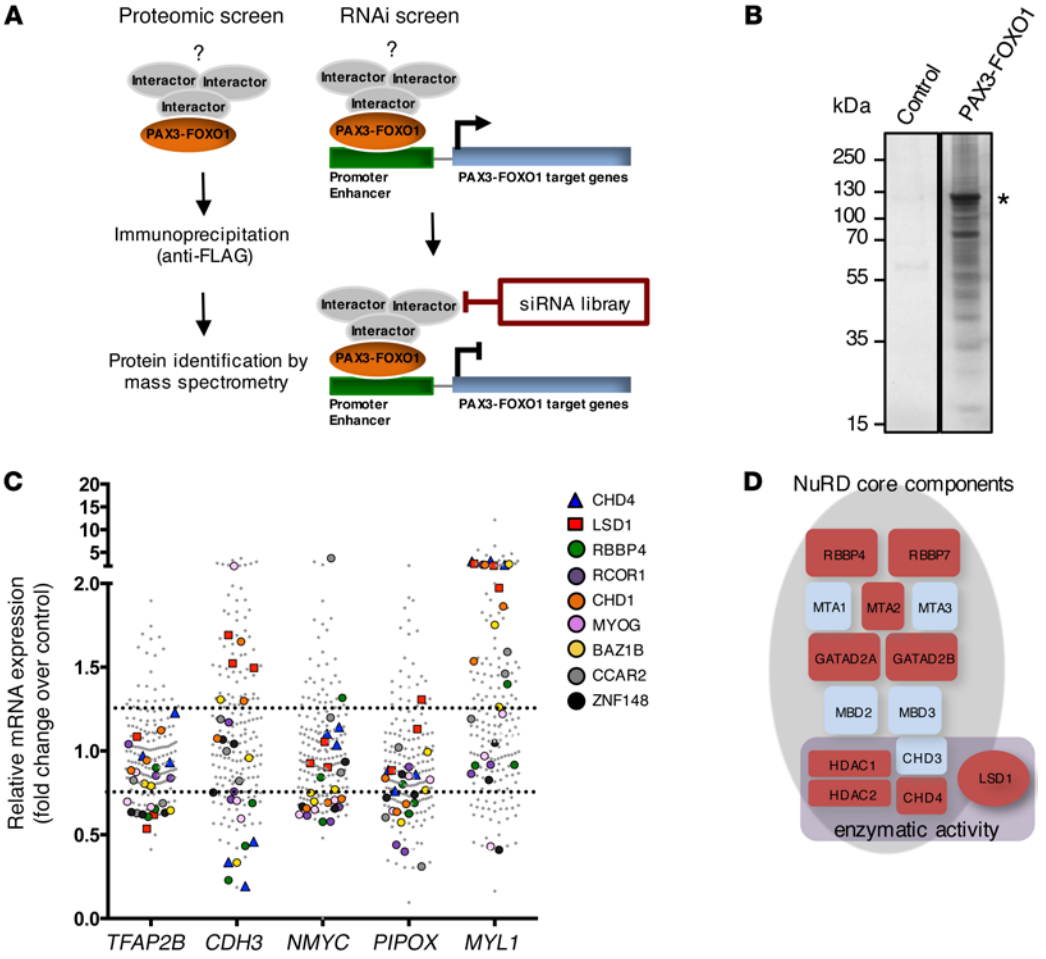

Figure 1. Two-step proteomic and siRNA screen identifies putative PAX3-FOXO1 interaction partners. (**A**) Schematic representation of the 2-step screening approach. (**B**) Representative silver-stained gel of FLAG immunoprecipitates from RMS13 cells transfected with FLAG-tagged PAX3-FOXO1. The asterisk marks the band corresponding to FLAG-PAX3-FOXO1. RMS13 cells expressing only FLAG-tag served as negative control. (**C**) SiRNA screening results for 60 candidate interactors. Each candidate was silenced using 3 different siRNAs per target. Fold change of expression of PAX3-FOXO1 target genes was measured by quantitative real-time PCR relative to RH4 cells treated with scrambled control. Values for the 9 candidate interactors resulting from the siRNA screen are indicated. (**D**) Schematic representation of the NuRD complex. General subunit composition and components with enzymatic activity are displayed. NuRD subunits identified by MS are marked in red.

or indirect target genes downstream of PAX3-FOXO1 (29–32). In contrast, epigenetic mechanisms that would orchestrate the network of aberrant transcriptional activity at the level of PAX3-FOXO1 itself are largely unknown.

Here, we used a 2-step screening approach to identify novel PAX3-FOXO1 coregulators that might be essential for its aberrant transcriptional activity. We identified chromodomain helicase DNA binding protein 4 (CHD4) as crucial coregulator of PAX3-FOXO1 activity that co-occupies a subset of PAX3-FOXO1 binding regions. Our results suggest that CHD4 represents a novel oncogenic epigenetic chromatin remodeler required to orchestrate a subsignature of PAX3-FOXO1 target genes and thereby maintain tumor cell survival.

Results

A 2-step screening approach identifies several NuRD complex components as PAX3-FOXO1 coregulators. To identify PAX3-FOXO1–interacting proteins in FP-RMS that may contribute to its aberrant activity and serve as rational therapeutic targets, we designed a 2-step screening approach (Figure 1A). First, we performed mass spectrometry (MS) analysis of proteins that interact with PAX3-FOXO1. FP-RMS cells RH4 and RMS13 were transiently transfected with triple FLAG-tagged PAX3-FOXO1 or, as control, FLAG only expression vectors. Using both N- and C-terminally FLAG-tagged PAX3-FOXO1 prevented loss of interactors due to interference with the localization of the tag. Additionally, we combined high-level CMV promoter–driven PAX3-FOXO1 expression for maximum protein recovery with low-level long terminal repeat promoter–driven PAX3-FOXO1 expression, minimizing nonspecific interactions (Supplemental Figure 1A; supplemental material available online with this article; doi:10.1172/JCI85057DS1). We analyzed 4 independent FLAG-affinity purifications of PAX3-FOXO1 together with controls (Figure 1B) by MS, which confirmed the presence of PAX3-FOXO1 with a minimum coverage of 54% in each replicate experiment including PAX3-FOXO1 breakpoint-specific peptides (Supplemental Figure 1B). Considering only proteins found with a

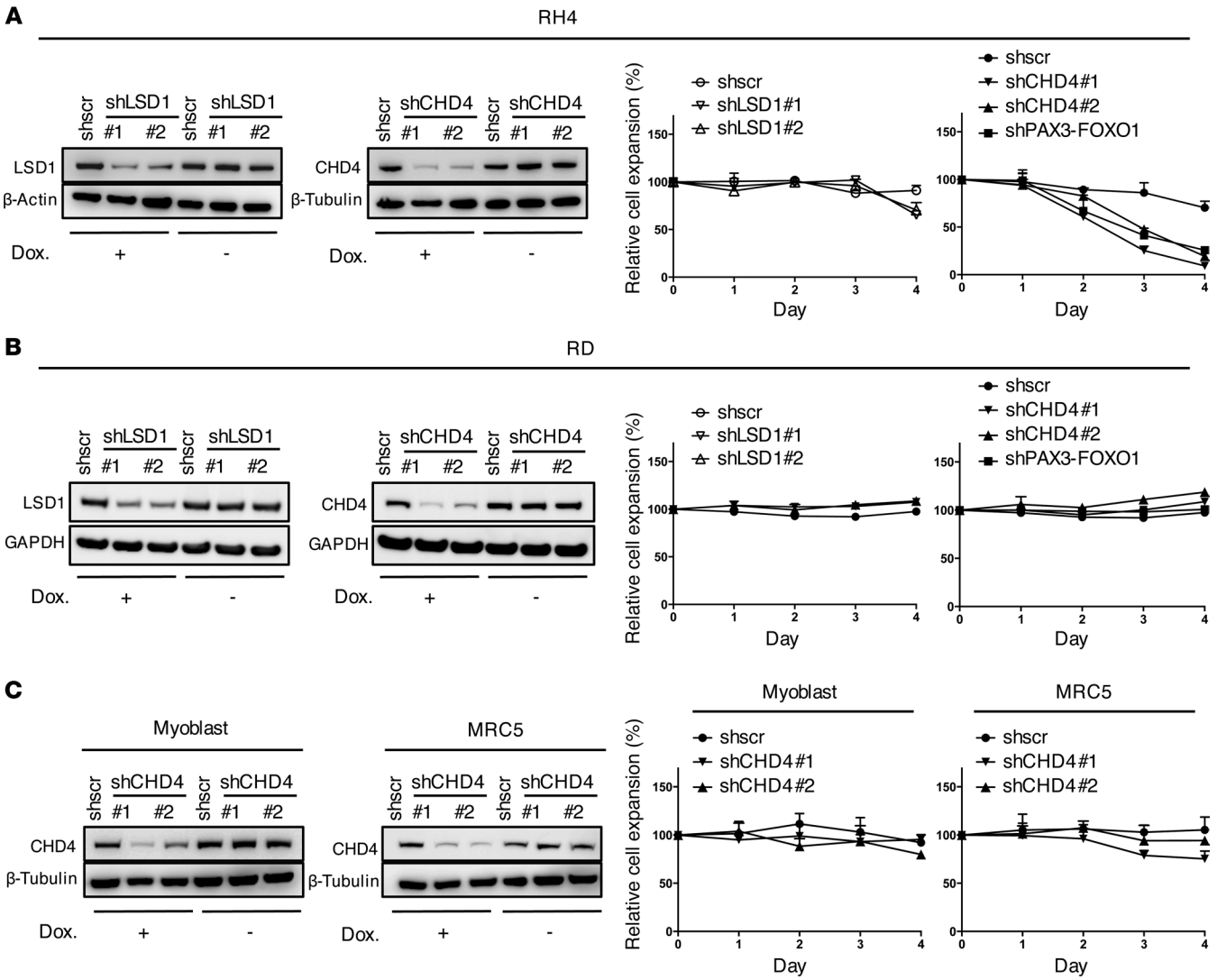

Figure 2. Silencing of CHD4 affects specifically FP-RMS cell expansion. (**A**) Left panels: Western blots after knockdown of CHD4 or LSD1 in RH4 cell extracts 72 hours after induction with doxycycline (Dox.). Uninduced cells served as negative control, and actin or tubulin was used as loading control. Right panels: RH4 cell expansion was measured by WST assay at indicated time points after silencing of LSD1 or CHD4 by shRNA induction with doxycycline. Data represent the mean ± SD of 3 independent experiments and were normalized to uninduced cells. PAX3-FOXO1 knockdown served as positive control. (**B**) Western blots and cell expansion after LSD1 or CHD4 knockdown in RD cells treated and analyzed as in **A**. (**C**) Western blots and cell proliferation after CHD4 knockdown in human myoblast or MRC5 cells treated and analyzed as in **A**.

minimum of 2 unique peptides in at least 2 replicate experiments and none of the controls, we identified 230 putative PAX3-FOXO1 candidate interactors (Supplemental Table 1).

To narrow down the list to candidate interactors that potentially would affect PAX3-FOXO1 transcriptional activity, we applied a subsequent siRNA screen (Figure 1A) targeting 60 preselected candidate interactors (Supplemental Table 2). Preselection was based on number of unique peptides identified by MS combined with possible function as described in the literature. Additionally, we used gene expression data from different FP-RMSs to include only candidates with detectable expression levels (24). RH4 cells, whose gene expression signature represents most closely tumor biopsies (19, 33), were individually transfected with 3 unique siRNA sequences per target and transcript levels of 5 PAX3-FOXO1 target genes measured after 48 hours. These genes represent different classes of genes affected by PAX3-FOXO1, including those activated from enhancer sites (*NMYC*), from intronic sites or promoters (*PIPOX* and *CDH3*), or via unknown mechanisms (*TFAP2B*), as well as 1 repressed gene (*MYL1*) (19, 34–37). For each siRNA, fold change of target gene expression relative to scrambled control was determined by quantitative real-time PCR. We considered a fold change below 0.75 for activated and above 1.25 for repressed target genes as threshold to narrow down the list of candidates (Figure 1C). Considering, in addition, candidates affecting at least 2 target genes by at least 2 independent siRNAs, we identified 9 interactors that might potentially modulate PAX3-FOXO1 activity (Figure 1C and Supplemental Figure 2). Each of the 9 candidate interactors affected expression of only a subset of the target genes measured, reflecting potentially diverse mechanisms by which the fusion protein regulates target gene expression.

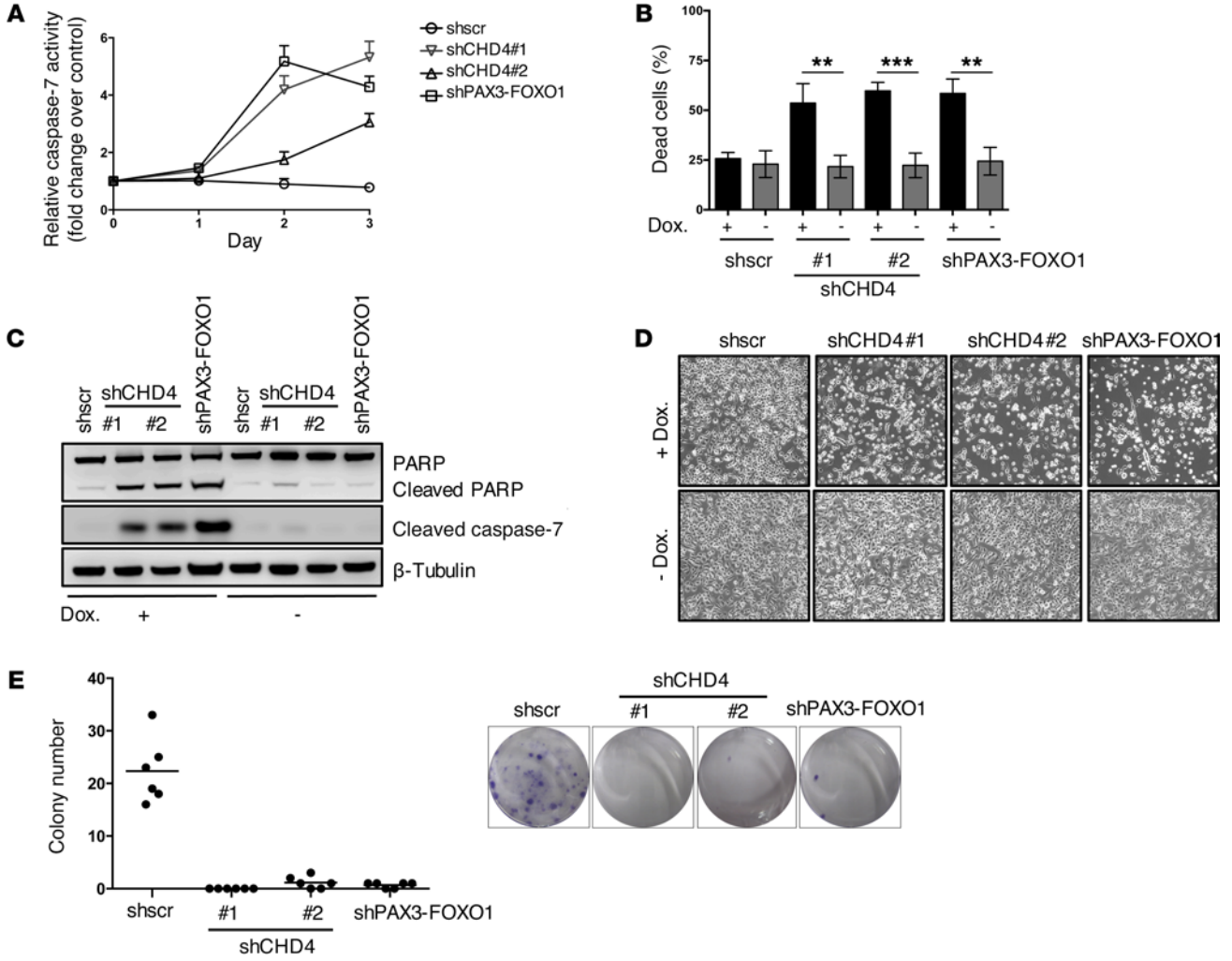

Figure 3. CHD4 knockdown decreases viability of FP-RMS cells. (**A**) Caspase-3/7 activity was measured in RH4 cells 72 hours after induction of CHD4 silencing. Fold change of caspase-3/7 activity was normalized to uninduced cells. Shscr-treated cells served as negative and PAX3-FOXO1 knockdown cells as positive controls. Values represent the mean ± SD of 3 independent biological experiments. (**B**) Percentage of dead cells 96 hours after induction of CHD4 silencing. Cells were stained with NucView Caspase-3 Substrate and 7AAD and analyzed by flow cytometry. Values represent the mean ± SD of 4 independent experiments (**$P < 0.01$; ***$P < 0.001$; uncorrected Fisher's LSD). (**C**) Western blot of PARP and cleaved caspase-7 from extracts of RH4 cells 72 hours after induction of CHD4 or PAX3-FOXO1 silencing and uninduced or shscr-treated control cells. (**D**) Representative phase-contrast images of RH4 cells 72 hours after induction of silencing with doxycycline (Dox.) transduced with indicated constructs; original magnification, ×100. (**E**) Clonogenic assays of RH4 cells 12 days after induction of CHD4 silencing. Quantitation of number of colonies with black lines representing the mean values. Representative images of crystal violet–stained colonies are shown in the right panel.

Remarkably, 5 of the 9 candidates represent components of chromatin remodeling complexes. CHD4, lysine-specific demethylase 1 (LSD1), and retinoblastoma binding protein 4 (RBBP4) are components of the nucleosome remodeling and deacetylase (NuRD) complex (38, 39), whereas REST corepressor 1 (RCOR1) and LSD1 can also be subunits of the CoREST complex (40, 41). In addition, the atypical tyrosine protein kinase BAZ1B is a bromodomain-containing protein found in different chromatin remodeling complexes (42). Importantly, our MS experiments identified several subunits of the NuRD complex, namely RBBP7 and histone deacetylases 1 and 2 (HDAC1 and HDAC2), which can function in diverse chromatin-regulating complexes, as well as GATA zinc finger–containing protein 2A (GATAD2A) and metastasis-associated protein 2 (MTA2), which are more specific components of the NuRD complex (Figure 1D and Supplemental Table 1).

In conclusion, our 2-step screening approach identified several chromatin remodeling proteins and complexes as potentially interacting directly or indirectly with PAX3-FOXO1, and points particularly to a specific role of the NuRD complex in regulating PAX3-FOXO1 target genes.

CHD4 is crucial for FP-RMS cell expansion in vitro. To further study the biological role of the NuRD complex, we selected CHD4 and LSD1 as members of the complex, which have not been studied in RMS and possess a potentially druggable enzymatic activity, and assessed whether their silencing would affect FP-RMS cell expansion. We established an inducible lentiviral expression system of stable cell lines expressing LSD1- or CHD4-specific shRNAs upon doxycycline treatment. Silencing was validated on mRNA and protein expression levels 72 hours after induction of shRNA expression with 2 independent

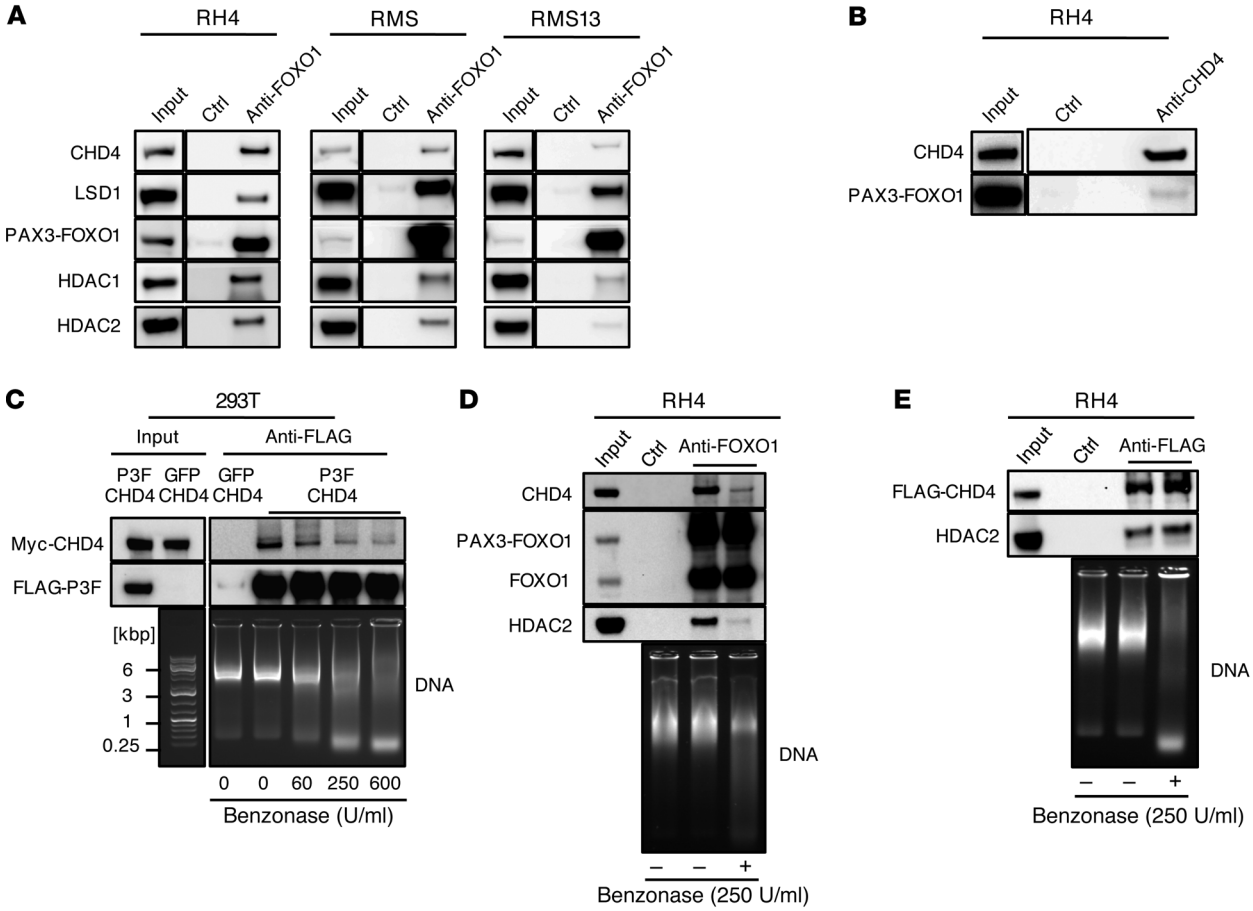

Figure 4. CHD4 and PAX3-FOXO1 mainly interact via short DNA fragments. (**A**) Representative Western blots of endogenous PAX3-FOXO1 immunoprecipitates and lysates from 3 different FP-RMS cell lines. PAX3-FOXO1 was immunoprecipitated by the anti-FOXO1 antibody, and uncoated beads served as negative control (ctrl). (**B**) Representative Western blot of reciprocal CHD4 immunoprecipitate from RH4 cells. Endogenous CHD4 was immunoprecipitated by an anti-CHD4 antibody, and uncoated beads served as negative control (ctrl). (**C**) Western blot detection of indicated proteins in anti-FLAG immunoprecipitates and lysates from 293T cells transfected with FLAG-tagged P3F and Myc-tagged CHD4. Cell lysate was treated with indicated amounts of Benzonase during immunoprecipitation to digest the DNA, and DNA digestion was evaluated by agarose gel electrophoresis. (**D**) Western blot detection of indicated endogenous proteins in anti-FOXO1 immunoprecipitates from RH4 cells. Uncoated beads served as negative control (ctrl). Lysates were digested or not with 250 U/ml Benzonase, and DNA digestion was evaluated by agarose gel electrophoresis. (**E**) Western blot detection of indicated endogenous proteins in anti-FLAG immunoprecipitates from RH4 cells with stable knock-in of 3X FLAG at C-terminus of CHD4. Uncoated beads served as negative control (Ctrl). Lysates were digested or not with 250 U/ml Benzonase, and DNA digestion was evaluated by agarose gel electrophoresis.

sequences per target in 2 different FP-RMS cell lines (RH4 and RMS) (Figure 2A and Supplemental Figure 3, A and B). As positive control we used a shRNA sequence based on a recently published PAX3-FOXO1 breakpoint-specific siRNA (siPF2) (ref. 18 and Supplemental Figure 3, A and B). Cell expansion was measured over 4 days and normalized to control cells to exclude doxycycline-mediated effects (43). We found that LSD1 depletion had only minor effects on RH4 expansion. In contrast, CHD4 knockdown decreased cell numbers of both cell lines to less than 30% of control, comparable to depletion of the fusion protein itself (Figure 2A and Supplemental Figure 3C). To assess specificity for FP-RMS, we performed the same experiments in the fusion-negative RMS (FN-RMS) cell line RD and did not observe any influence of CHD4 depletion on cell numbers (Figure 2B). Importantly, neither nontransformed human fibroblasts nor normal primary myoblast cells were affected by CHD4 depletion (Figure 2C). Taken together these results demonstrate that CHD4 but not LSD1 function is essential for expansion specifically of translocation-driven FP-RMS cells, whereas it does not affect the growth of normal cells. Based on this result, we focused our studies on CHD4.

CHD4 promotes FP-RMS cell survival. Next, we asked whether reduction in cell numbers upon CHD4 depletion was due to loss of viability. To this end, we measured caspase-3/7 activity after induction of CHD4 silencing in 2 FP-RMS cell lines (RH4 and RMS) over a period of 3 days (Figure 3A and Supplemental Figure 4A). Both CHD4-specific shRNAs increased caspase-3/7 activity between 3- and 6-fold in RH4 cells and to a lesser extent in the FP-RMS cell line RMS relative to control. Increased caspase-3/7 activity after CHD4 knockdown was additionally confirmed in both cell lines by detection of active caspase-3/7 in individual cells by flow cytometry (Supplemental Figure 5). Furthermore, we quantified the percentage of dead cells 96 hours after induction of CHD4 silencing by combined detection of

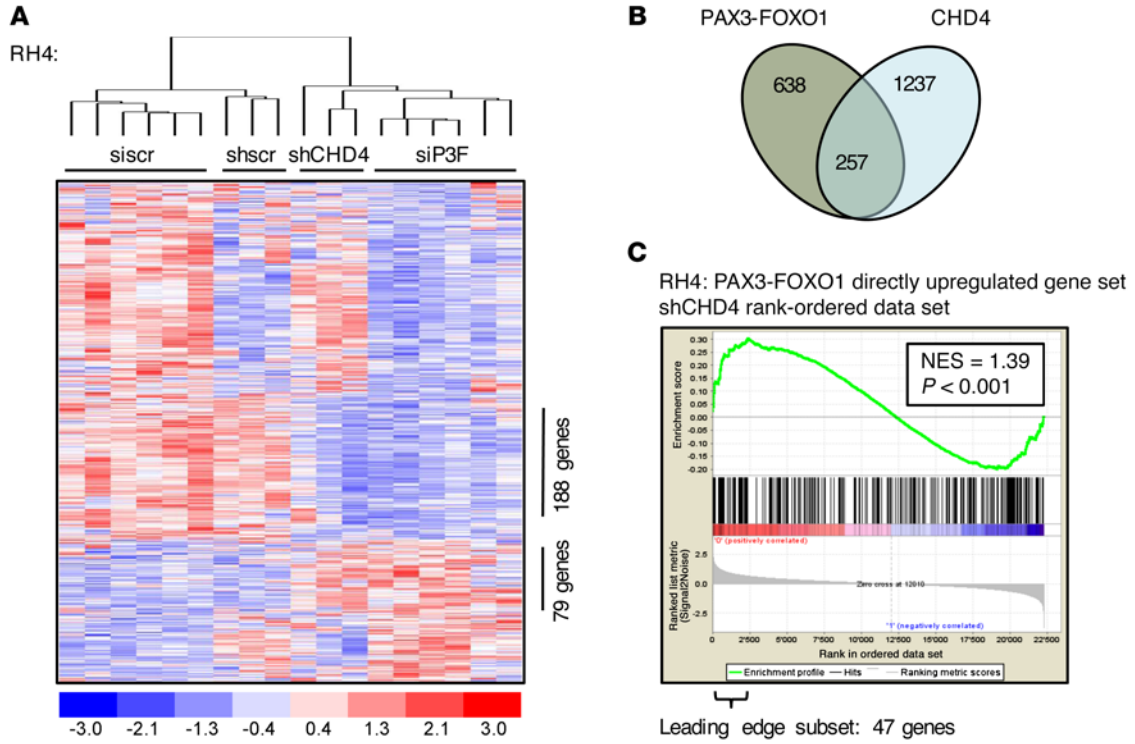

Figure 5. CHD4 coregulates PAX3-FOXO1–activated target genes. (**A**) Heat map of unsupervised hierarchical clustering of PAX3-FOXO1 (siP3F) and CHD4 (shCHD4#1) knockdown profiles using the gene set directly regulated by PAX3-FOXO1 (638 genes differently expressed between siP3F-treated and siscr-treated [24 and 48 hours after transfection] or untreated RH4 cells; fold change >1.5; $P < 0.05$). Each column represents a different time point for shRNA-treated cells in singlicate and for siRNA-treated cells in duplicate (24, 48, 72 hours from left to right for each condition). Numbers of coexpressed genes between these transcriptional profiles are displayed. (**B**) Schematic representation of genes directly regulated by PAX3-FOXO1 and CHD4 (fold change >1.5; $P < 0.05$ between siP3F-treated and siscr-treated [24 and 48 hours after transfection] or untreated RH4 cells and between shCHD4-expressing and shscr-expressing RH4 cells [24 and 48 hours after doxycycline induction] or uninduced control cells). The coregulated subset as read out from the heatmap clustering in **A** is displayed. (**C**) GSEA using CHD4-regulated genes in RH4 cells as the rank-ordered data set and targets directly upregulated by PAX3-FOXO1 as the gene set (259 genes differently expressed between siP3F-treated and siscr-treated or untreated RH4 cells 24 and 48 hours after transfection; fold change >1.7; $P < 0.05$). Normalized enrichment score (NES) and P value are shown.

active caspase-3/7–positive and 7-aminoactinomycin D–positive (7-AAD–positive) cells (Figure 3B, Supplemental Figure 4B, and Supplemental Figure 6). This revealed that CHD4 depletion increased the percentage of dead cells from 22% to 48% in RH4 (21% to 51% in the cell line RMS), a level similar to that seen with depletion of PAX3-FOXO1 (increase from 20% to 43% in RH4, 21% to 47% in the cell line RMS), whereas caspase-3/7 was active in the majority of dying cells. In addition, protein analysis of CHD4-depleted cells demonstrated increased levels of cleaved caspase-7 and cleaved PARP in FP-RMS cells (Figure 3C and Supplemental Figure 4C), which were not detected in normal human fibroblasts or myoblasts (data not shown). These markers of apoptotic cell death also correlated with an increased number of floating cells, observed only in FP-RMS cells but not in normal human fibroblasts or myoblasts (Figure 3D, Supplemental Figure 4D, and Supplemental Figure 7). Finally, CHD4 depletion impaired the ability of both FP-RMS cell lines to form colonies in clonogenic cell survival assays (Figure 3E and Supplemental Figure 4E). In summary, the data demonstrate that the antiproliferative effect of CHD4 depletion in FP-RMS cells is due to increased cell death, suggesting that CHD4 expression is necessary to promote tumor cell survival.

CHD4 interacts with PAX3-FOXO1 via short DNA fragments. To get mechanistic insights into the interaction of PAX3-FOXO1 with NuRD complex members, we first immunoprecipitated endogenous PAX3-FOXO1 from 3 different FP-RMS cell lines (RH4, RMS13, and RMS) with an anti-FOXO1 antibody and found specific coprecipitation of HDAC1, HDAC2, LSD1, and CHD4, but not with the negative control (Figure 4A). Importantly, these interactions persisted also after silencing of WT FOXO1 with a FOXO1-specific siRNA (Supplemental Figure 8), indicating that they do not depend on WT FOXO1, which can also be immunoprecipitated by the anti-FOXO1 antibody. In support, immunoprecipitation of CHD4 was able to pull down PAX3-FOXO1 (Figure 4B) in the reciprocal approach. To further investigate whether interaction of CHD4 and PAX3-FOXO1 occurs by direct protein-protein contact, we expressed tagged proteins together in HEK293T cells and immunoprecipitated FLAG-PAX3-FOXO1. Interestingly, under normal conditions we observed efficient pull-down of CHD4, whereas digestion of DNA with increasing amounts of Benzonase gradually decreased coprecipitated CHD4 (Figure 4C). Similar observations were made with the endogenous proteins CHD4 and HDAC2 in FP-RMS cells (Figure 4D), whereas the previously characterized

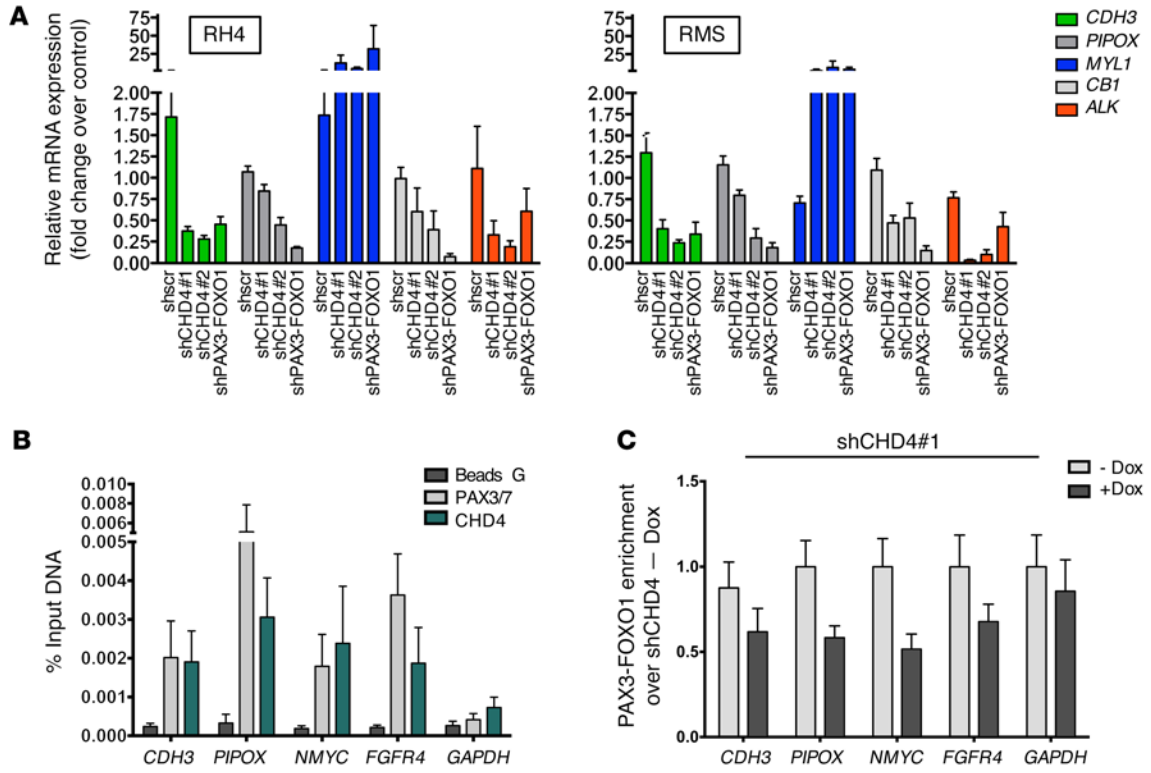

Figure 6. CHD4 colocalizes to PAX3-FOXO1 binding sites. (**A**) Expression levels of indicated PAX3-FOXO1 target genes were quantified by quantitative real-time PCR after CHD4 knockdown in 2 FP-RMS cell lines. Bar charts are geometric means from 4 independent experiments with 95% CI ($P < 0.05$; Dunnett's multiple comparison test). Fold change of mRNA expression was normalized to uninduced cells with PAX3-FOXO1 knockdown serving as positive control. (**B**) ChIP was performed in RH4 cells using PAX3/7 and CHD4 antibodies on known PAX3-FOXO1 DNA binding sites in target genes, and only beads without antibody served as negative control (Beads G). Bar charts indicate the mean ± SEM of at least 3 independent biological replicates. Abundance of precipitated DNA fragments was measured by quantitative PCR, and results are presented as the percentage of input. The *GAPDH* promoter region served as negative control. (**C**) Effect on PAX3-FOXO1 binding upon CHD4 silencing (48 hours of incubation with doxycycline) was determined by ChIP as described in **B** and is shown as fold enrichment over PAX3-FOXO1 ChIP signal in the presence of CHD4 (no doxycycline control).

interaction between the NuRD components CHD4 and HDAC2 was unaffected by Benzonase treatment (Figure 4E). The data suggest that the interaction between PAX3-FOXO1 and CHD4 is mainly dependent on short DNA fragments to which the 2 factors bind in close proximity.

CHD4 depletion affects PAX3-FOXO1–activated target genes. Given the importance of CHD4 for FP-RMS cell survival and the observation that none of the initial candidates, including CHD4, affected all 5 PAX3-FOXO1 target genes in our primary siRNA screen, we sought to explore the global effects of CHD4 depletion on the PAX3-FOXO1 transcriptome. Therefore, we performed gene expression profiling on RNA isolated from RH4 cells 24, 48, and 72 hours after CHD4 depletion. We analyzed gene expression changes in regard to a recently published PAX3-FOXO1 gene expression signature (19), since we hypothesized that predominantly direct PAX3-FOXO1 target genes might be affected, focusing on genes showing at least 1.5-fold up- or downregulation at the early time points (24 and 48 hours) after PAX3-FOXO1 silencing. Hierarchical clustering of this target gene signature revealed that the CHD4 gene expression signature clusters together with the PAX3-FOXO1 signature and not with the control-treated cells (Figure 5A). Comparison of both signatures affirms that CHD4 depletion affects only a subset of the PAX3-FOXO1 target gene signature and mainly overlaps in genes that are normally upregulated by PAX3-FOXO1. Indeed, the overlapping coregulated subset is composed of 257 genes of a total of 638 genes affected by PAX3-FOXO1 depletion (Figure 5B). This further supports a model in which PAX3-FOXO1 drives expression of target gene subsets in concert with different (epigenetic) mechanisms and indicates that CHD4 is important for a subset of PAX3-FOXO1–mediated transcriptionally activated target genes. Gene set enrichment analysis (GSEA) confirmed this conclusion by unraveling a significant correlation of CHD4 expression signature only with genes directly activated by PAX3-FOXO1, not with genes inhibited by PAX3-FOXO1 nor with genes showing changes only at later time points (Figure 5C and Supplemental Figure 9). The leading-edge subset of this coactivated gene set consists of 47 genes (Supplemental Table 3). Besides the PAX3-FOXO1 target genes already examined in our primary siRNA screen, we found that CHD4 inhibition had a strong impact on expression of anaplastic lymphoma kinase (ALK) and cannabinoid receptor 1 (CB1), both PAX3-FOXO1 target genes that have been implicated in the tumor biology of FP-RMS (44, 45). In particular, expression of CB1 has been correlated with increased invasiveness of FP-RMS tumors (45). Therefore, our data suggest that CHD4 expression is required to coregulate a subset of PAX3-FOXO1 target genes that are necessary and sufficient to promote FP-RMS survival.

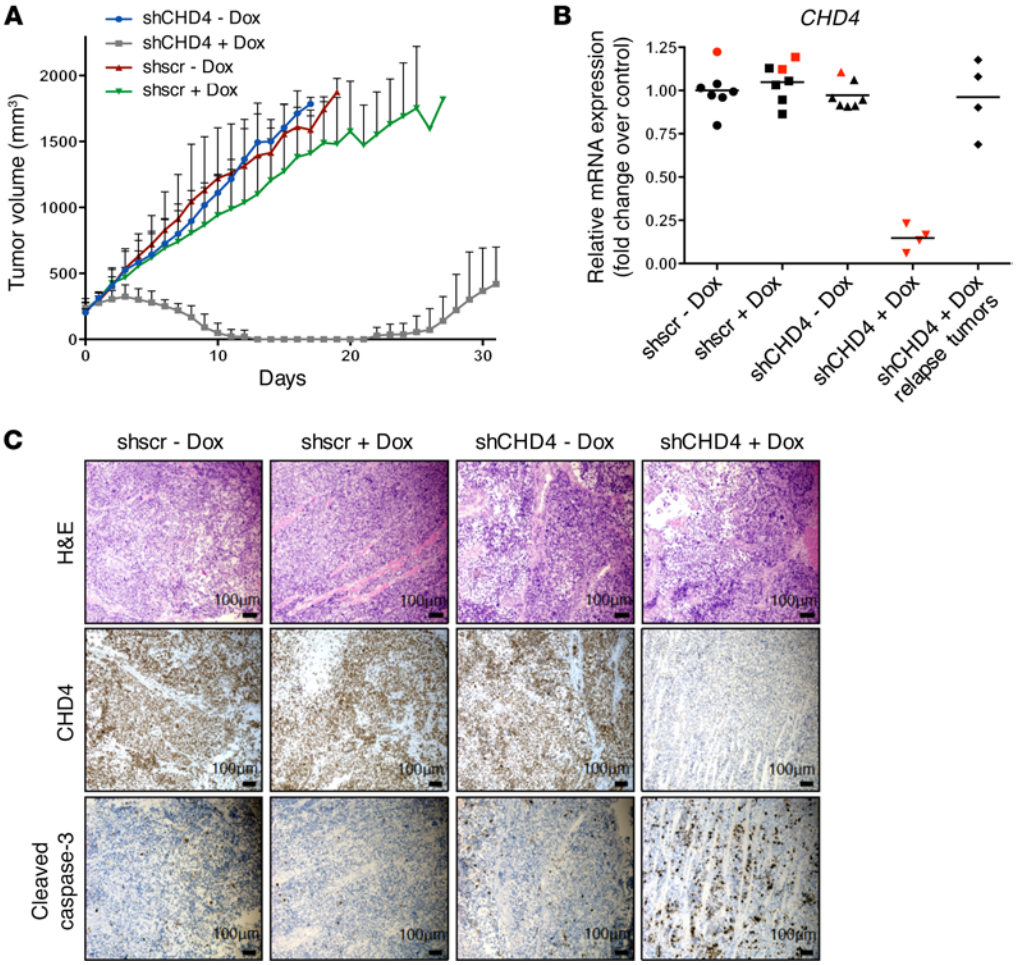

Figure 7. CHD4 inhibition causes tumor regression in mouse xenografts. In vivo treatment of NOD/SCID mice engrafted with RH4 cells containing stably integrated doxycycline-inducible shCHD4#1 or shscr expression vectors. Mice bearing palpable tumors were treated i.p. for 2 days with either vehicle control or doxycycline at a dose of 53.3 mg/kg. Additionally, they were fed with doxycycline-supplemented or control food. (**A**) Absolute tumor volumes of FP-RMS xenografts as measured by caliper. Vehicle-treated control groups consisted of 6 mice, and doxycycline-treated groups consisted of 5 (shscr) and 7 (shCHD4#1) mice. (**B**) CHD4 mRNA expression level in indicated tumors of doxycycline- or vehicle control–treated mice analyzed by quantitative real-time PCR. Tumors were isolated at the end of treatment (black; tumor set shown in **A**) or after 4 days of treatment (red; separate tumor set). Fold changes of mRNA expression are normalized to the mean of CHD4 expression in shscr vehicle control–treated tumors. (**C**) Immunohistochemical staining of tumors with indicated antibodies. Tumors were isolated 4 days after the start of treatment.

CHD4 and PAX3-FOXO1 colocalize at PAX3-FOXO1 target genes. To examine whether the fusion protein indeed binds to target genes together with CHD4, we first validated selected PAX3-FOXO1 target genes with a known role in FP-RMS pathology, in 2 different FP-RMS cell lines (RH4 and RMS) 72 hours after induction of CHD4 depletion (Figure 6A) (22, 36, 44–46). We observed similar modulation of target gene levels by depleting CHD4 or PAX3-FOXO1. Next, we used ChIP on previously published PAX3-FOXO1 binding sites at selected target genes (34). It is of note that the sites of PAX3-FOXO1 binding are mostly enhancer regions (e.g., *NMYC*, *FGFR4*, *PIPOX*) and only rarely promoter regions (e.g., *CDH3*) promoter region). We analyzed the occupancy of CHD4 around PAX3-FOXO1 binding sites together with the *GAPDH* promoter region as negative control. We were able to measure ChIP PCR values that were low, but significantly above background. First, we confirmed PAX3-FOXO1 binding to all selected target genes, but not to the *GAPDH* promoter (Figure 6B). For precipitation of PAX3-FOXO1 we used a PAX3/7 antibody that has been recently used in ChIP analysis, since expression of WT PAX3 is 1,000-fold lower and, as such, negligible in RH4 cells (19, 36). Next, we observed that CHD4 also is specifically bound to the same fragments of all target genes. Interestingly, prior depletion of CHD4 using the stable shRNA cells caused a clear reduction of PAX3-FOXO1 binding on all target genes (Figure 6C), whereas addition of doxycycline had no effect on the binding of PAX3-FOXO1 in the *GAPDH* promoter region. This indicates that fusion protein binding is modulated by the presence of CHD4. Further, CHD4 and PAX3-FOXO1 co-occupy the same sites in PAX3-FOXO1 target genes, suggesting coregulation of target gene expression by both CHD4 and the fusion protein, at least for a subset of target genes.

CHD4 inhibition causes tumor regression in mouse xenografts. Based on the observation that CHD4 inhibition reduces FP-RMS cell proliferation and survival in vitro, we aimed to assess whether CHD4 inhibition affects FP-RMS tumor growth in vivo using xenograft mouse models. Immunocompromised NOD/SCID mice were engrafted s.c. with RH4 cells, containing stably integrated tet-inducible shCHD4#1 or shscr expression vectors. Mice with established tumors of varying size (170–400 mm^3) were treated i.p. with either vehicle or doxycycline solution at a dose of 53.3 mg/kg on 2 consecutive days. Additionally, starting from the first day of treatment, mice were fed with doxycycline-supplemented or control food. Strikingly, we observed regression of tumor size upon CHD4 depletion, even in tumors with a large starting volume of 380 mm^3 (Figure 7A). Both immunohistochemical analyses of tumors and determination of mRNA levels showed strong reduction of CHD4 expression in all depleted tumors (Figure 7, B and C). The CHD4-depleted tumors also displayed high expression of the apoptotic marker cleaved caspase-3, similar to our previous in vitro findings (Figure 7C).

Interestingly, tumors that recurred after 21 days of treatment (Figure 7A) showed recovery of CHD4 mRNA expression levels to nearly 100% and comparable to those in untreated controls (Figure 7B). Hence, it is likely that cells with insufficient CHD4 depletion did finally outgrow.

To investigate whether CHD4 depletion also affects FN-RMS tumors, we measured RD cell tumor growth in vivo under doxycycline addition. Interestingly, we observed only a slight growth retardation by doxycycline per se, but no tumor regression upon CHD4 depletion (Supplemental Figure 10). This was not due to generally lower CHD4 expression levels in FN-RMS, as both subtypes have a similar expression pattern assessed on a human RMS tumor tissue microarray (Supplemental Figure 11). Taken together, these data clearly demonstrate that the epigenetic regulator CHD4 plays an important role in FP-RMS tumor progression and survival, and identify CHD4 as novel potential treatment target for this pediatric sarcoma.

Discussion

Although PAX3-FOXO1 expression on its own might not be sufficient for RMS development, it is the dominant oncogenic driver and sole relevant genetic abnormality in a large fraction of tumors (27). Since direct pharmacological inhibition of transcription factors is still elusive, blocking the activity of PAX3-FOXO1 accessory proteins may therefore represent a viable alternative therapeutic strategy.

In the current study, we provide the first comprehensive analysis of the interactome of PAX3-FOXO1 directly purified from FP-RMS cells. A biochemical proteomic approach was used in a first step and then followed by a functional siRNA screen evaluating expression levels of selected PAX3-FOXO1 target genes as readout (47). Interestingly, the final list of 9 putative accessory proteins revealed 6 protein subunits of chromatin remodeling complexes, suggesting that the fusion protein influences diverse epigenetic mechanisms. This is remarkably similar to the function of other fusion proteins in pediatric sarcomas, such as the Ewing sarcoma fusion EWS-FLI1, which is known to use divergent chromatin remodeling mechanisms to orchestrate target gene expression (48), and the SSX-SS18 fusion, which disrupts the BAF complex in synovial sarcoma (49).

Most of the chromatin remodelers identified are subunits of the NuRD complex, namely CHD4, LSD1, and RBBP4. Further, we identified additional components of this complex in our initial proteomic screen, namely RBBP7, MTA2, GATAD2A, GATAD2B, HDAC1, and HDAC2, suggesting that PAX3-FOXO1 might indeed interact with the NuRD complex. It has been demonstrated that this remodeling complex is also required for EWS-FLI1–mediated transcriptional repression (50), and inhibition of LSD1 using small-molecule inhibitors impaired Ewing sarcoma cell viability (51). Hence, we next thought to identify components of the complex that might be necessary for PAX3-FOXO1 function. Whereas we observed only minor effects on proliferation of FP-RMS cells upon genetic depletion of LSD1, CHD4 depletion decreased FP-RMS cell proliferation and induced cell death to an extent similar to that seen with inhibition of PAX3-FOXO1 itself, without influencing proliferation of normal human fibroblasts and myoblasts, thereby opening a potential therapeutic window. Additionally, experimental evidence supports a role for CHD4 specifically in FP-RMS cell survival, which is a core function of the fusion protein. Interestingly, when comparing gene signatures after CHD4 depletion with the transcriptional profile regulated directly by PAX3-FOXO1 (33, 46, 52), we identified around 40% of target genes as being coregulated, suggesting that genes responsible for cell survival might be part of this new limited subset signature. The exact contribution of the individual genes to cell survival remains to be determined, as it is not to be expected that they all act as drivers. However, it is important to emphasize that few of the well-characterized PAX3-FOXO1 target genes were identified in our gene set enrichment analysis (Supplemental Table 3). These included *ALK*, *CDH3*, and *TFAP2B*, of which only *TFAP2B* has been implicated in cell survival previously (19). Attempts to correlate expression of these genes individually with patient survival data using published expression profiling data sets have failed so far (data not shown). Nevertheless, our data suggest that other epigenetic regulators such as members of the SWI/SNF complex that were also identified in the MS analysis might be involved in modulating expression of the remaining subset of PAX3-FOXO1 target genes.

Mechanistically, it is, at this stage, not possible to determine the exact composition of the complex at the different locations. Although we first confirmed individual interactions for CHD4, LSD1, HDAC1, and HDAC2 under standard conditions, it remained unclear whether they all bind to the fusion protein via direct protein-protein interaction. Hence, we carried out immunoprecipitation experiments under conditions where potential residual DNA is digested. Indeed, this dramatically reduced the amount of coprecipitated PAX3-FOXO1, allowing us to conclude that they bind nearby to small DNA fragments, and unlikely via direct protein-protein interactions. Whether this also applies to the other chromatin complexes identified remains to be demonstrated.

Given the strong impact of CHD4 on PAX3-FOXO1 target gene activation, it might even be plausible that CHD4 acts in a NuRD complex–independent manner. This hypothesis is supported by the observation that depletion of several other NuRD subunits by siRNA (HDAC1, HDAC2, RBBP7, GATAD2A, MTA2) had only minor effects on target gene expression (data not shown). Interestingly, it has been suggested previously that CHD4 can function outside of the repressive NuRD complex to activate transcription (53–56). Such oncogenic properties of CHD4 in FP-RMS are in agreement with findings in other tumors such as glioblastoma and colorectal cancer (57, 58). On the other hand, CHD4 is frequently mutated in endometrial cancers and colorectal or gastric cancers with microsatellite instability (59, 60), suggesting that its oncogenic or tumor-suppressive functions depend on the cellular context.

Finally, we demonstrate that CHD4 depletion induces regression of FP-RMS tumors in xenograft mouse models, but, importantly, had no effect on FN-RMS tumors. Altogether, this validates CHD4 inhibition as a potential targeted strategy to inhibit an essential coregulator of PAX3-FOXO1 activity. Nevertheless, at present we cannot exclude the possibility that additional mechanisms besides modulation of PAX3-FOXO1 target genes contribute to the growth-suppressive phenotype, given that CHD4 has been also implicated in DNA damage response and G1/S cell cycle progression (61–64).

To our knowledge our experiments identify the first chromatin remodeler contributing to PAX3-FOXO1 transcriptional regulation. Since CHD4 contains ATPase activity, it might be an amenable target to inhibit PAX3-FOXO1 activity indirectly. However, the ATPase domain is highly conserved within the *SNF2* family of helicase-like ATPases and therefore might be difficult to target specifically (65). Alternatively, CHD4 contains 2 plant homeodomain (PHD) fingers and tandem chromodomains required for ATP-dependent chromatin remodeling, which could provide targets for drug design similar to the striking success in targeting BET bromodomains (66–70). Given the recognized role of the NuRD complex in other sarcomas, such as Ewing sarcoma, inhibition of CHD4 might be of therapeutic benefit for a broader range of tumors as well.

Methods

Cell lines. The aRMS cell lines RH4 (provided by Peter Houghton, Greehey Children's Cancer Research Institute, San Antonio, Texas, USA), RMS13 (provided by Roland Kappler, Ludwig–Maximilian University, Munich, Germany), and RMS (provided by Janet Shipley, The Institute of Cancer Research, London, United Kingdom), and MRC5 normal lung fibroblast cells, the eRMS cell line RD, and the HEK293T cell line for the production of lentiviral particles (all 3 purchased from the ATCC, LGC Promochem, Molsheim, France), were routinely maintained in DMEM (Sigma-Aldrich) supplemented with 10% FBS (Thermo Fisher Scientific, LuBioScience), 2 mM L-glutamine, and 100 U/ml penicillin/streptomycin. The human myoblast cell line KM155C25Dist (provided by Vincent Mouly, Institut de Myologie, Paris, France) was maintained in Skeletal Muscle Cell Growth Medium Supplement Mix (PromoCell), 15% FBS, 2 mM L-glutamine, and 100 U/ml penicillin/streptomycin. All cells were cultured in 5% CO_2 at 37°C. All RMS and the MRC5 cell line were authenticated by short tandem repeat analysis (STR) profiling in 2014/2015 and positively matched (71). As the human myoblast cell line has not yet been characterized by STR profiling, negative matching with all available cell lines in the database was used for verification.

Purification of FLAG-PAX3-FOXO1. RH4 and RMS13 cells were transfected with pCMV-N/C-FLAG-PAX3-FOXO1 or pBABE puro N-FLAG-PAX3-FOXO1 in 15-cm plates and lysed 48 hours after transfection with a mild lysis buffer containing 50 mM Tris-Cl (pH 7.5), 125 mM NaCl, 10% glycerol, 0.3% NP-40, 1.5 mM $MgCl_2$, 25 mM NaF, 10 mM sodium β-glycerolphosphate, 5 mM sodium pyrophosphate, and 2 mM sodium orthovanadate and supplemented with Complete Mini Protease Inhibitor cocktail (Roche Diagnostics). FLAG-PAX3-FOXO1 was immunoprecipitated using 100 μl Dynabeads Protein G (10004D, 30 mg/ml; Thermo Fisher Scientific, LuBioScience) per plate coupled to 8 μg monoclonal ANTI-FLAG M2 antibody (F1804, Sigma-Aldrich). After incubation for 10 minutes at 4°C, beads were washed 3 times with lysis buffer, and bound FLAG-PAX3-FOXO1 was eluted with 40 μl elution buffer containing 200 μg/ml 3X FLAG peptide (F4799, Sigma-Aldrich), 50 mM Tris-Cl (pH 7.4), and 150 mM NaCl. For each MS experiment the eluates of 12 plates (RMS13) or 24 plates (RH4) were combined and concentrated using the Uppa-Protein Concentrate kit (786-120, GBiosciences, VWR International AG), according to the manufacturer's instructions. After addition of 4X NuPAGE LDS buffer (Thermo Fisher Scientific, LuBioScience), proteins were separated by gel electrophoresis and stained with colloidal Coomassie (InstantBlue, Expedeon) or by silver staining (Bio-Rad). Coomassie-stained gels were cut into slices and prepared for MS as described in Supplemental Methods.

SiRNA screen. A siRNA library (Supplemental Table 2) targeting 60 candidate interactors identified in the primary MS analysis (Ambion Silencer Select Custom siRNA Library, Thermo Fisher Scientific, LuBioScience) was used with 3 unique siRNAs per target. Silencer Select Negative Control siRNA (4390846) served as nontargeting negative control and PAX3-FOXO1 breakpoint-specific siRNA (18) as positive control (both Ambion, Thermo Fisher Scientific, LuBioScience). For silencing of gene expression, 40,000 RH4 cells were reverse transfected with 4.6 nM siRNA in 24-well format using Interferin transfection reagent according to the manufacturer's protocol (Polyplus Transfection SA). RNA was harvested 48 hours after transfection and used for quantitative real-time PCR.

Western blotting. For Western blots, total protein extracts were obtained from cells lysed with RIPA buffer containing 50 mM Tris-Cl (pH 7.5), 150 mM NaCl, 1% NP-40, 0.5% Na-deoxycholate, 1 mM EGTA, 0.1% SDS, 50 mM NaF, 10 mM sodium β-glycerolphosphate, 5 mM sodium pyrophosphate, and 1 mM sodium orthovanadate and supplemented with Complete Mini Protease Inhibitor cocktail (Roche Diagnostics). Protein concentration was measured with Pierce BCA protein Assay Kit (Thermo Fisher Scientific, LuBioScience). Proteins were separated using 4%–12% Bis-Tris SDS-PAGE gels (Thermo Fisher Scientific, LuBioScience) and transferred to nitrocellulose membranes (Protran, Schleicher & Schuell). After blocking with 5% milk powder in TBS/0.1% Tween, membranes were incubated with primary antibodies overnight at 4°C. After washing in TBS/0.1% Tween, membranes were incubated with HRP-linked IgG antibodies. Proteins were detected by chemiluminescence using ECL detection reagent or SuperSignal West Femto Maximum Sensitivity Substrate (both Thermo Fisher Scientific) after washing in TBS/0.1% Tween.

Coimmunoprecipitation. Cells were lysed in 1 ml lysis buffer per 10-cm dish and incubated for 1 hour at 4°C with antibody directed against the protein of interest coupled to Dynabeads Protein G (Novex by Thermo Fisher Scientific, LuBioScience) or empty beads as negative control. Antibodies used were directed against FOXO1 (C-20) (sc-9808, Santa Cruz Biotechnology), FLAG-tag (Sigma-Aldrich), or CHD4 (A301-082A, polyclonal, rabbit, Bethyl Laboratories). Benzonase (Novagen) was added to the lysate during this incubation when indicated. After washing 4 times with lysis buffer, proteins were eluted with 200 μg/ml 3X FLAG peptide or with 1X NuPAGE LDS sample buffer (Thermo Fisher Scientific, LuBioScience) at 70°C and analyzed by Western blotting.

Detection of DNA fragmentation by agarose gel electrophoresis. Twelve microliters of immunoprecipitation supernatants were run on 1% agarose gels, and resolved DNA was stained with GelRed.

Cell viability assay. For cell viability assays, aRMS, MRC5, and human myoblast cells were cultured in 96-well format. At various time points after induction of shRNA expression, cell viability was measured by WST-1 assay (Roche Diagnostics) and normalized to uninduced control cells. At least 3 biological replicates were performed for each experiment.

Clonogenic assay. For clonogenic assays, 24 hours after induction of shRNA expression, 2,000 (RH4) or 5,500 (RMS) cells were seeded on 6-well plates. Doxycycline-containing complete medium was changed every second to third day. After 12 days of incubation, cells were fixed with 4% formaldehyde solution and stained with 0.05% crystal violet, and colonies were counted.

Cell death and caspase-3/7 activity assays. Cells were seeded in white 384-well plates with clear bottom (Greiner Bio-One). Caspase activity was determined by Caspase-Glo 3/7 Assay (Promega) at indicated time points according to the manufacturer's instructions. Luminescence was measured using the multidetection microplate reader Synergy HT (Bio-Tek Instruments). For analysis of caspase-3/7 activity in individual cells, they were stained with NucView 405 Caspase-3 Substrate (Biotium, Chemie Brunschwig AG) and analyzed by flow cytometry. For quantification of dead cells, NucView 405 Caspase-3 Substrate was combined with 7-AAD staining solution (BD Biosciences).

Microarray analysis. For gene expression profiling, RNA of RH4 cells was isolated from untreated cells, and 24, 48, and 72 hours after induction of CHD4 silencing (shCHD4#1) using RNeasy Kit (Qiagen), with DNase digestion step, according to the manufacturer's instructions. Affymetrix GeneChip expression analysis was performed using HGU133plus2 arrays (Affymetrix Inc.). cRNA target synthesis and experimental procedures for GeneChip hybridization and scanning were carried out according to the "GeneChip eukaryotic small sample target labeling technical note" (Affymetrix, ATLAS Biolabs). Expression data were combined with previous published PAX3-FOXO1 gene expression data (19). The data sets are available for unrestricted download from Gene Expression Omnibus under the accession numbers GSE73480 and GSE73483. Batch normalization and unsupervised hierarchical clustering were performed using dChip software. Gene set enrichment analysis (GSEA) was performed using GSEA version 2 software (Broad Institute) (72).

ChIP. ChIP has been recently described (73). For each ChIP assay, 30 μg of cross-linked and sonicated chromatin was incubated with 4 μg of antibody. Purified ChIP DNA was measured for enrichment of PAX3-FOXO1 binding sites by quantitative real-time PCR using TaqMan assays. Primer set and TaqMan probes are described in Supplemental Table 5.

Xenograft studies. For the generation of xenograft, 5×10^6 RH4 cells, containing stably integrated tetracycline-inducible shCHD4#1 or shscr expression construct, were transplanted s.c. into 8- to 12-week-old female NOD/SCID mice (Charles River). After first detection of palpable tumors, mice were treated i.p. with either sterile PBS or doxycycline solution at 53.3 mg/kg on 2 consecutive days. Doxycycline was diluted in PBS at a concentration of 8 mg/ml. Additionally, starting from the first day of treatment, mice were fed with doxycycline-supplemented food (625 mg doxycycline/kg) or control food. Tumor size was determined every day by measurement of 2 diameters (d_1, d_2) in right angles using a digital caliper. Total tumor volumes were calculated by the formula $V = (4/3)\pi r^3$; $r = (d_1 + d_2)/4$. Control mice were euthanized when reaching a tumor volume of 1,700 mm^3.

Immunohistochemistry. Tumor tissue was cut into slices of 4–5 mm and fixed in 4% vol/vol neutralized formalin for 24 hours at room temperature, processed in a Pathos Delta (Milestone) overnight with ethanol and isopropanol for dehydration and paraffin embedding. Samples were molded in paraffin blocks and sectioned on a Leica rotary microtome into 3-μm-thick slices. After drying at 64°C for 30 minutes, slides were stained with H&E in a Leica Autostainer using Mayers HTX (catalog 01820, Histolab; 10 minutes) and Eosin aqueous (catalog 2C-140, Waldeck; 0.5%, 5 minutes). After staining, slides were dehydrated and mounted with Pertex. Immunohistochemistry was performed on Leica BondMax instruments using Refine HRP Kits (Leica DS9800) including all buffer solutions from Leica Microsystems, processed according to the manufacturer's guidelines. Antigen retrieval for Cleaved Caspase-3 (9661, Cell Signaling Technology; dilution 1:500) was performed with Epitope Retrieval buffer (Leica AR9640) at 100°C for 60 minutes. Slides for CHD4 ChIP Grade (ab70469, Abcam Ltd.; dilution 1:500) were pretreated with Epitope Retrieval buffer (Leica AR9961) at 100°C for 30 minutes.

Statistics. All statistical analysis (with the exception of microarray and GSEA analysis) was performed with GraphPad Prism software version 6. One-way ANOVA tests with Dunnett's method for multiple comparisons with a 95% CI were used to analyze data involving 2 or more test groups and a control group. All statistical analysis of quantitative real-time PCR data was performed on the level of the $\Delta\Delta C_t$ values. One-way ANOVA tests with uncorrected Fisher's LSD method with a 95% CI were used for pairwise comparisons of several test groups. *P* values less than 0.05 were considered statistically significant. Data are presented as the geometric mean with 95% CI or as mean ± SD unless otherwise noted.

Study approval. All animal experiments were approved by the Swiss veterinary authorities and were performed according to the animal license 208/2012.

Information regarding plasmids, transfection methods, lentiviral transduction, immunohistochemistry, tissue array, antibodies, quantitative real-time PCR, MS, and database searching including protein identification can be found in Supplemental Methods.

Author contributions

MB, MW, JM, and BWS performed the conceptual work and designed the experiments. MB, MW, JM, NS, and DL performed the experiments and, together with BWS, analyzed the data. PN aided MS data acquisition and interpretation. KM provided the myoblast cells. RS aided ChIP–quantitative real-time PCR data acquisition and interpretation. MB and BWS wrote the paper.

Acknowledgments

We thank Silvia Behnke (Sophistolab AG) and the computer scientists at the Functional Genomics Center Zurich for their excellent service. We also deeply acknowledge Sergio Leone and Damian Dalcher (University of Zurich) for helpful discussions and their support. This work was supported by the Swiss National Science Fund (31003A-138460), Swiss Research Foundation Child and Cancer, Foundation Empiris, the Cancer League of the Canton of Zurich, and Forschungskredit University of Zurich.

Address correspondence to: Beat W. Schäfer, Department of Oncology, Children's Hospital Zurich, Steinwiesstrasse 75, 8032 Zurich, Switzerland. Phone: 41.44.266.7553; E-mail: Beat.Schaefer@kispi.uzh.ch.

1. Breneman JC, et al. Prognostic factors and clinical outcomes in children and adolescents with metastatic rhabdomyosarcoma — a report from the Intergroup Rhabdomyosarcoma Study IV. *J Clin Oncol*. 2003;21(1):78–84.
2. Van Gaal JC, De Bont ES, Kaal SE, Versleijen-Jonkers Y, van der Graaf WT. Building the bridge between rhabdomyosarcoma in children, adolescents and young adults: the road ahead. *Crit Rev Oncol Hematol*. 2012;82(3):259–279.
3. Barr FG. Gene fusions involving PAX and FOX family members in alveolar rhabdomyosarcoma.

Oncogene. 2001;20(40):5736–5746.
4. Skapek SX, et al. PAX-FOXO1 fusion status drives unfavorable outcome for children with rhabdomyosarcoma: a children's oncology group report. *Pediatr Blood Cancer*. 2013;60(9):1411–1417.
5. Missiaglia E, et al. Genomic imbalances in rhabdomyosarcoma cell lines affect expression of genes frequently altered in primary tumors: an approach to identify candidate genes involved in tumor development. *Genes Chromosomes Cancer*. 2009;48(6):455–467.
6. Duan F, et al. Genomic and clinical analysis of fusion gene amplification in rhabdomyosarcoma: a report from the Children's Oncology Group. *Genes Chromosomes Cancer*. 2012;51(7):662–674.
7. Sorensen PH, et al. PAX3-FKHR and PAX7-FKHR gene fusions are prognostic indicators in alveolar rhabdomyosarcoma: a report from the children's oncology group. *J Clin Oncol*. 2002;20(11):2672–2679.
8. Davis RJ, Barr FG. Fusion genes resulting from alternative chromosomal translocations are overexpressed by gene-specific mechanisms in alveolar rhabdomyosarcoma. *Proc Natl Acad Sci U S A*. 1997;94(15):8047–8051.
9. Miller PJ, Hollenbach AD. The oncogenic fusion protein Pax3-FKHR has a greater post-translational stability relative to Pax3 during early myogenesis. *Biochim Biophys Acta*. 2007;1770(10):1450–1458.
10. del Peso L, González VM, Hernández R, Barr FG, Núñez G. Regulation of the forkhead transcription factor FKHR, but not the PAX3-FKHR fusion protein, by the serine/threonine kinase Akt. *Oncogene*. 1999;18(51):7328–7333.
11. Fredericks WJ, et al. The PAX3-FKHR fusion protein created by the t(2;13) translocation in alveolar rhabdomyosarcomas is a more potent transcriptional activator than PAX3. *Mol Cell Biol*. 1995;15(3):1522–1535.
12. Keller C, Arenkiel BR, Coffin CM, El-Bardeesy N, DePinho RA, Capecchi MR. Alveolar rhabdomyosarcomas in conditional Pax3:Fkhr mice: cooperativity of Ink4a/ARF and Trp53 loss of function. *Genes Dev*. 2004;18(21):2614–2626.
13. Khan J, et al. cDNA microarrays detect activation of a myogenic transcription program by the PAX3-FKHR fusion oncogene. *Proc Natl Acad Sci U S A*. 1999;96(23):13264–13269.
14. Naini S, et al. Defining the cooperative genetic changes that temporally drive alveolar rhabdomyosarcoma. *Cancer Res*. 2008;68(23):9583–9588.
15. Ren YX, et al. Mouse mesenchymal stem cells expressing PAX-FKHR form alveolar rhabdomyosarcomas by cooperating with secondary mutations. *Cancer Res*. 2008;68(16):6587–6597.
16. Scheidler S, Fredericks WJ, Rauscher FJ, Barr FG, Vogt PK. The hybrid PAX3-FKHR fusion protein of alveolar rhabdomyosarcoma transforms fibroblasts in culture. *Proc Natl Acad Sci U S A*. 1996;93(18):9805–9809.
17. Scuoppo C, et al. The oncogenic transcription factor PAX3-FKHR can convert fibroblasts into contractile myotubes. *Exp Cell Res*. 2007;313(11):2308–2317.
18. Kikuchi K, et al. Effects of PAX3-FKHR on malignant phenotypes in alveolar rhabdomyosarcoma. *Biochem Biophys Res Commun*. 2008;365(3):568–574.
19. Ebauer M, Wachtel M, Niggli FK, Schäfer BW. Comparative expression profiling identifies an in vivo target gene signature with TFAP2B as a mediator of the survival function of PAX3/FKHR. *Oncogene*. 2007;26(51):7267–7281.
20. Bernasconi M, Remppis A, Fredericks WJ, Rauscher FJ, Schäfer BW. Induction of apoptosis in rhabdomyosarcoma cells through down-regulation of PAX proteins. *Proc Natl Acad Sci U S A*. 1996;93(23):13164–13169.
21. Fredericks WJ, Ayyanathan K, Herlyn M, Friedman JR, Rauscher FJ. An engineered PAX3-KRAB transcriptional repressor inhibits the malignant phenotype of alveolar rhabdomyosarcoma cells harboring the endogenous PAX3-FKHR oncogene. *Mol Cell Biol*. 2000;20(14):5019–5031.
22. Marshall AD, Grosveld GC. Alveolar rhabdomyosarcoma – The molecular drivers of PAX3/7-FOXO1-induced tumorigenesis. *Skelet Muscle*. 2012;2(1):25.
23. Zhu B, Davie JK. New insights into signalling-pathway alterations in rhabdomyosarcoma. *Br J Cancer*. 2015;112(2):227–231.
24. Wachtel M, Rakic J, Okoniewski M, Bode P, Niggli F, Schäfer BW. FGFR4 signaling couples to Bim and not Bmf to discriminate subsets of alveolar rhabdomyosarcoma cells. *Int J Cancer*. 2014;135(7):1543–1552.
25. Cao L, et al. Addiction to elevated insulin-like growth factor I receptor and initial modulation of the AKT pathway define the responsiveness of rhabdomyosarcoma to the targeting antibody. *Cancer Res*. 2008;68(19):8039–8048.
26. Vogelstein B, Papadopoulos N, Velculescu VE, Zhou S, Diaz LA, Kinzler KW. Cancer genome landscapes. *Science*. 2013;339(6127):1546–1558.
27. Shern JF, et al. Comprehensive genomic analysis of rhabdomyosarcoma reveals a landscape of alterations affecting a common genetic axis in fusion-positive and fusion-negative tumors. *Cancer Discov*. 2014;4(2):216–231.
28. Esteller M. Epigenetics in cancer. *N Engl J Med*. 2008;358(11):1148–1159.
29. Ciarapica R, et al. The Polycomb group (PcG) protein EZH2 supports the survival of PAX3-FOXO1 alveolar rhabdomyosarcoma by repressing FBXO32 (Atrogin1/MAFbx). *Oncogene*. 2014;33(32):4173–4184.
30. Ciarapica R, et al. Deregulated expression of miR-26a and Ezh2 in rhabdomyosarcoma. *Cell Cycle*. 2009;8(1):172–175.
31. Walters ZS, et al. JARID2 is a direct target of the PAX3-FOXO1 fusion protein and inhibits myogenic differentiation of rhabdomyosarcoma cells. *Oncogene*. 2014;33(9):1148–1157.
32. Lee MH, Jothi M, Gudkov AV, Mal AK. Histone methyltransferase KMT1A restrains entry of alveolar rhabdomyosarcoma cells into a myogenic differentiated state. *Cancer Res*. 2011;71(11):3921–3931.
33. Wachtel M, et al. Gene expression signatures identify rhabdomyosarcoma subtypes and detect a novel t(2;2)(q35;p23) translocation fusing PAX3 to NCOA1. *Cancer Res*. 2004;64(16):5539–5545.
34. Cao L, et al. Genome-wide identification of PAX3-FKHR binding sites in rhabdomyosarcoma reveals candidate target genes important for development and cancer. *Cancer Res*. 2010;70(16):6497–6508.
35. Mercado GE, et al. Identification of PAX3-FKHR-regulated genes differentially expressed between alveolar and embryonal rhabdomyosarcoma: focus on MYCN as a biologically relevant target. *Genes Chromosomes Cancer*. 2008;47(6):510–520.
36. Thuault S, et al. P-cadherin is a direct PAX3-FOXO1A target involved in alveolar rhabdomyosarcoma aggressiveness. *Oncogene*. 2013;32(15):1876–1887.
37. De Pittà C, et al. Gene expression profiling identifies potential relevant genes in alveolar rhabdomyosarcoma pathogenesis and discriminates PAX3-FKHR positive and negative tumors. *Int J Cancer*. 2006;118(11):2772–2781.
38. Wang Y, et al. LSD1 is a subunit of the NuRD complex and targets the metastasis programs in breast cancer. *Cell*. 2009;138(4):660–672.
39. Allen HF, Wade PA, Kutateladze TG. The NuRD architecture. *Cell Mol Life Sci*. 2013;70(19):3513–3524.
40. Shi YJ, Matson C, Lan F, Iwase S, Baba T, Shi Y. Regulation of LSD1 histone demethylase activity by its associated factors. *Mol Cell*. 2005;19(6):857–864.
41. You A, Tong JK, Grozinger CM, Schreiber SL. CoREST is an integral component of the CoREST- human histone deacetylase complex. *Proc Natl Acad Sci U S A*. 2001;98(4):1454–1458.
42. Barnett C, Krebs JE. WSTF does it all: a multifunctional protein in transcription, repair, and replication. *Biochem Cell Biol*. 2011;89(1):12–23.
43. Ahler E, et al. Doxycycline alters metabolism and proliferation of human cell lines. *PLoS One*. 2013;8(5):e64561.
44. van Gaal JC, et al. Anaplastic lymphoma kinase aberrations in rhabdomyosarcoma: clinical and prognostic implications. *J Clin Oncol*. 2012;30(3):308–315.
45. Marshall AD, Lagutina I, Grosveld GC. PAX3-FOXO1 induces cannabinoid receptor 1 to enhance cell invasion and metastasis. *Cancer Res*. 2011;71(24):7471–7480.
46. Williamson D, et al. Fusion gene-negative alveolar rhabdomyosarcoma is clinically and molecularly indistinguishable from embryonal rhabdomyosarcoma. *J Clin Oncol*. 2010;28(13):2151–2158.
47. Thalhammer V, et al. PLK1 phosphorylates PAX3-FOXO1, the inhibition of which triggers regression of alveolar Rhabdomyosarcoma. *Cancer Res*. 2015;75(1):98–110.
48. Riggi N, et al. EWS-FLI1 utilizes divergent chromatin remodeling mechanisms to directly activate or repress enhancer elements in Ewing sarcoma. *Cancer Cell*. 2014;26(5):668–681.
49. Kadoch C, Crabtree GR. Reversible disruption of mSWI/SNF (BAF) complexes by the SS18-SSX oncogenic fusion in synovial sarcoma. *Cell*. 2013;153(1):71–85.
50. Sankar S, et al. Mechanism and relevance of EWS/FLI-mediated transcriptional repression in Ewing sarcoma. *Oncogene*. 2013;32(42):5089–5100.
51. Sankar S, et al. Reversible LSD1 inhibition interferes with global EWS/ETS transcriptional activity and impedes Ewing sarcoma tumor growth. *Clin Cancer Res*. 2014;20(17):4584–4597.
52. Davicioni E, Finckenstein FG, Shahbazian V,

Buckley JD, Triche TJ, Anderson MJ. Identification of a PAX-FKHR gene expression signature that defines molecular classes and determines the prognosis of alveolar rhabdomyosarcomas. *Cancer Res*. 2006;66(14):6936–6946.
53. Reynolds N, O'Shaughnessy A, Hendrich B. Transcriptional repressors: multifaceted regulators of gene expression. *Development*. 2013;140(3):505–512.
54. Williams CJ, et al. The chromatin remodeler Mi-2β is required for CD4 expression and T cell development. *Immunity*. 2004;20(6):719–733.
55. Amaya M, et al. Mi2β-mediated silencing of the fetal γ-globin gene in adult erythroid cells. *Blood*. 2013;121(17):3493–3501.
56. Murawska M, Hassler M, Renkawitz-Pohl R, Ladurner A, Brehm A. Stress-induced PARP activation mediates recruitment of Drosophila Mi-2 to promote heat shock gene expression. *PLoS Genet*. 2011;7(7):e1002206.
57. Cai Y, et al. The NuRD complex cooperates with DNMTs to maintain silencing of key colorectal tumor suppressor genes. *Oncogene*. 2014;33(17):2157–2168.
58. Chudnovsky Y, et al. ZFHX4 interacts with the NuRD core member CHD4 and regulates the glioblastoma tumor-initiating cell state. *Cell Rep*. 2014;6(2):313–324.
59. Zhao S, et al. Landscape of somatic single-nucleotide and copy-number mutations in uterine serous carcinoma. *Proc Natl Acad Sci U S A*. 2013;110(8):2916–2921.
60. Kim MS, Chung NG, Kang MR, Yoo NJ, Lee SH. Genetic and expressional alterations of CHD genes in gastric and colorectal cancers. *Histopathology*. 2011;58(5):660–668.
61. O'Shaughnessy A, Hendrich B. CHD4 in the DNA-damage response and cell cycle progression: not so NuRDy now. *Biochem Soc Trans*. 2013;41(3):777–782.
62. Polo SE, Kaidi A, Baskcomb L, Galanty Y, Jackson SP. Regulation of DNA-damage responses and cell-cycle progression by the chromatin remodelling factor CHD4. *EMBO J*. 2010;29(18):3130–3139.
63. Larsen DH, et al. The chromatin-remodeling factor CHD4 coordinates signaling and repair after DNA damage. *J Cell Biol*. 2010;190(5):731–740.
64. Smeenk G, Wiegant WW, Vrolijk H, Solari AP, Pastink A, van Attikum H. The NuRD chromatin-remodeling complex regulates signaling and repair of DNA damage. *J Cell Biol*. 2010;190(5):741–749.
65. Flaus A, Martin DM, Barton GJ, Owen-Hughes T. Identification of multiple distinct Snf2 subfamilies with conserved structural motifs. *Nucleic Acids Res*. 2006;34(10):2887–2905.
66. Mansfield RE, et al. Plant homeodomain (PHD) fingers of CHD4 are histone H3-binding modules with preference for unmodified H3K4 and methylated H3K9. *J Biol Chem*. 2011;286(13):11779–11791.
67. Musselman CA, et al. Binding of the CHD4 PHD2 finger to histone H3 is modulated by covalent modifications. *Biochem J*. 2009;423(2):179–187.
68. Filippakopoulos P, et al. Selective inhibition of BET bromodomains. *Nature*. 2010;468(7327):1067–1073.
69. Morra R, Lee BM, Shaw H, Tuma R, Mancini EJ. Concerted action of the PHD, chromo and motor domains regulates the human chromatin remodelling ATPase CHD4. *FEBS Lett*. 2012;586(16):2513–2521.
70. Bouazoune K, et al. The dMi-2 chromodomains are DNA binding modules important for ATP-dependent nucleosome mobilization. *EMBO J*. 2002;21(10):2430–2440.
71. Hinson AR, Jones R, Crose LE, Belyea BC, Barr FG, Linardic CM. Human rhabdomyosarcoma cell lines for rhabdomyosarcoma research: utility and pitfalls. *Front Oncol*. 2013;3:183.
72. Subramanian A, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A*. 2005;102(43):15545–15550.
73. Savić N, et al. lncRNA maturation to initiate heterochromatin formation in the nucleolus is required for exit from pluripotency in ESCs. *Cell Stem Cell*. 2014;15(6):720–734.

Manuscript 4

PAX3-FOXO1 establishes myogenic super enhancers and confers BET bromodomain vulnerability

Berkley E. Gryder[1], Marielle E. Yohe[1,2], Hsien-Chao Chou[1], Xiaohu Zhang[3], Joana G. Marques[4], Marco Wachtel[4], Beat Schaefer[4], Nirmalya Sen[1], Young Song[1], Alberto Gualtieri[5], Silvia Pomella[5], Rossella Rota[5], Abigail Cleveland[1], Xinyu Wen[1], Sivasish Sindiri[1], Jun S. Wei[1], Frederic G. Barr[6], Sudipto Das[7], Thorkell Andresson[7], Rajarshi Guha[3], Madhu Lal-Nag[3], Marc Ferrer[3], Jack F. Shern[1,2], Keji Zhao[8], Craig J. Thomas[3], and Javed Khan[1]

[1]Genetics Branch, NCI, NIH, Bethesda, MD, USA
[2]Pediatric Oncology Branch, CCR, NCI, NIH, Bethesda, MD, USA
[3]Division of Preclinical Innovation, NCATS, NIH, Rockville, MD, USA
[4]University Children's Hospital, Zurich, Switzerland
[5]Department of Oncohematology, Ospedale Pediatrico Bambino Gesù Research Institute, Viale San Paolo 15, 00146 Rome, Italy
[6]Laboratory of Pathology, NCI, NIH, Bethesda, MD, USA
[7]Laboratory of Proteomics and Analytical Technologies, ATC, NCI, Frederick, Maryland, USA
[8]Systems Biology Center, NHLBI, NIH, Bethesda, MD, USA

Conflict of Interest: The authors declare no conflict of interest.

Published manuscript.

Corresponding Author: Javed Khan, M.D.

Genetics Branch, NCI, NIH, Bethesda, MD, USA
khanjav@mail.nih.gov

Personal contribution: I performed several of the ChIP-seq assays used in the Figures 1C/D, 4A-E, S1C, S3C-E and S5A/F. Also, I contributed to the intellectual content of the paper.

Published OnlineFirst April 26, 2017; DOI: 10.1158/2159-8290.CD-16-1297

RESEARCH ARTICLE

PAX3–FOXO1 Establishes Myogenic Super Enhancers and Confers BET Bromodomain Vulnerability

Berkley E. Gryder[1], Marielle E. Yohe[1,2], Hsien-Chao Chou[1], Xiaohu Zhang[3], Joana Marques[4], Marco Wachtel[4], Beat Schaefer[4], Nirmalya Sen[1], Young Song[1], Alberto Gualtieri[5], Silvia Pomella[5], Rossella Rota[5], Abigail Cleveland[1], Xinyu Wen[1], Sivasish Sindiri[1], Jun S. Wei[1], Frederic G. Barr[6], Sudipto Das[7], Thorkell Andresson[7], Rajarshi Guha[3], Madhu Lal-Nag[3], Marc Ferrer[3], Jack F. Shern[1,2], Keji Zhao[8], Craig J. Thomas[3], and Javed Khan[1]

Published OnlineFirst April 26, 2017; DOI: 10.1158/2159-8290.CD-16-1297

ABSTRACT

Alveolar rhabdomyosarcoma is a life-threatening myogenic cancer of children and adolescent young adults, driven primarily by the chimeric transcription factor PAX3–FOXO1. The mechanisms by which PAX3–FOXO1 dysregulates chromatin are unknown. We find PAX3–FOXO1 reprograms the *cis*-regulatory landscape by inducing *de novo* super enhancers. PAX3–FOXO1 uses super enhancers to set up autoregulatory loops in collaboration with the master transcription factors MYOG, MYOD, and MYCN. This myogenic super enhancer circuitry is consistent across cell lines and primary tumors. Cells harboring the fusion gene are selectively sensitive to small-molecule inhibition of protein targets induced by, or bound to, PAX3–FOXO1-occupied super enhancers. Furthermore, PAX3–FOXO1 recruits and requires the BET bromodomain protein BRD4 to function at super enhancers, resulting in a complete dependence on BRD4 and a significant susceptibility to BRD inhibition. These results yield insights into the epigenetic functions of PAX3–FOXO1 and reveal a specific vulnerability that can be exploited for precision therapy.

SIGNIFICANCE: PAX3–FOXO1 drives pediatric fusion-positive rhabdomyosarcoma, and its chromatin-level functions are critical to understanding its oncogenic activity. We find that PAX3–FOXO1 establishes a myoblastic super enhancer landscape and creates a profound subtype-unique dependence on BET bromodomains, the inhibition of which ablates PAX3–FOXO1 function, providing a mechanistic rationale for exploring BET inhibitors for patients bearing PAX-fusion rhabdomyosarcoma. *Cancer Discov; 7(8); 884–99. ©2017 AACR.*

INTRODUCTION

Transcription factors (TF) recognize specific noncoding sequences across the genome, recruiting epigenetic machinery to regulate key cell identity genes, and are sequentially exchanged during development and differentiation (1). Oncogenic fusion genes involving TFs are predicted to profoundly alter normal developmental progression and cell identity in many malignancies (2).

Rhabdomyosarcoma (RMS) is a cancer of childhood and adolescence characterized by its inability to exit the proliferative myoblast-like state. Genomic and transcriptomic characterization implicates either chromosomal translocation resulting in the oncogenic fusion transcription factor PAX3/7–FOXO1 (fusion-positive alveolar subtype, FP-RMS) or mutations in receptor tyrosine kinase/RAS pathways (fusion-negative embryonal subtype, FN-RMS; refs. 3, 4). FP-RMS is characterized by a strikingly low somatic mutational burden indicating that the fusion gene is the primary oncogenic driver. Many other transcription/chromatin factor fusion gene–driven sarcomas have similar low mutational burdens (5–7). Importantly, patients with RMS who harbor a PAX3 fusion are more likely to be metastatic at presentation, relapse despite aggressive therapy, and have very poor survival (8), underscoring the critical need to develop therapeutic strategies for this subset of patients.

Early (PAX family TFs) and late (MYOG) regulators of normal myogenesis are temporally mutually exclusive in normal muscle development, yet FP-RMS tumors concurrently express high levels of PAX3–FOXO1 and the myogenic MYOD, MYOG, as well as MYCN (9, 10). Although the transcriptional perturbation caused by the PAX fusions has been previously reported (11, 12), the chromatin mechanisms by which PAX fusions dysregulate the myogenic program are unknown. In this work, we interrogated the underlying epigenetics that enforce the myogenic and oncogenic transcriptional program of cell lines and clinical tumor samples with PAX3 fusions. In charting the genome-wide landscape of histone modifications, we discovered that PAX3–FOXO1 drives expression of its target oncogenes by creating large deposits of active histone marks exclusively at enhancers, collaboratively with myogenic TFs, and by recruiting chromatin reader bromodomain-containing protein 4 (BRD4), which function at looped enhancer–promoter pairs within topological domain boundaries. Integrating epigenetic and mechanistic drug screening data exposed multiple biological nodes of chemical vulnerability, including BET bromodomains. The BRD4 inhibitor JQ1 has recently shown efficacy in RMS (13), but no mechanistic connection has been shown between BRD4 and PAX3–FOXO1. Here, we report that BRD4 inhibition disrupts a hitherto undiscovered PAX3–FOXO1 interaction with BRD4, causes a rapid degradation of the fusion gene, and ablates its transcriptional output, thus revealing a subtype-selective therapeutic vulnerability to BRD4 inhibition.

[1]Genetics Branch, NCI, NIH, Bethesda, Maryland. [2]Pediatric Oncology Branch, Center for Cancer Research, NCI, NIH, Bethesda, Maryland. [3]Division of Preclinical Innovation, National Center for Advancing Translational Sciences, NIH, Rockville, Maryland. [4]University Children's Hospital, Zurich, Switzerland. [5]Department of Oncohematology, Ospedale Pediatrico Bambino Gesù Research Institute, Rome, Italy. [6]Laboratory of Pathology, NCI, NIH, Bethesda, Maryland. [7]Laboratory of Proteomics and Analytical Technologies, Advanced Technologies Center, NCI, Frederick, Maryland. [8]Systems Biology Center, National Heart, Lung, and Blood Institute, NIH, Bethesda, Maryland.

Note: Supplementary data for this article are available at Cancer Discovery Online (http://cancerdiscovery.aacrjournals.org/).

Corresponding Author: Javed Khan, NIH, 37 Convent Drive, Building 37, Room 2016B, Bethesda, MD 20892. Phone: 240-760-6135; Fax: 301-480-0314; E-mail: khanjav@mail.nih.gov

doi: 10.1158/2159-8290.CD-16-1297

Published OnlineFirst April 26, 2017; DOI: 10.1158/2159-8290.CD-16-1297

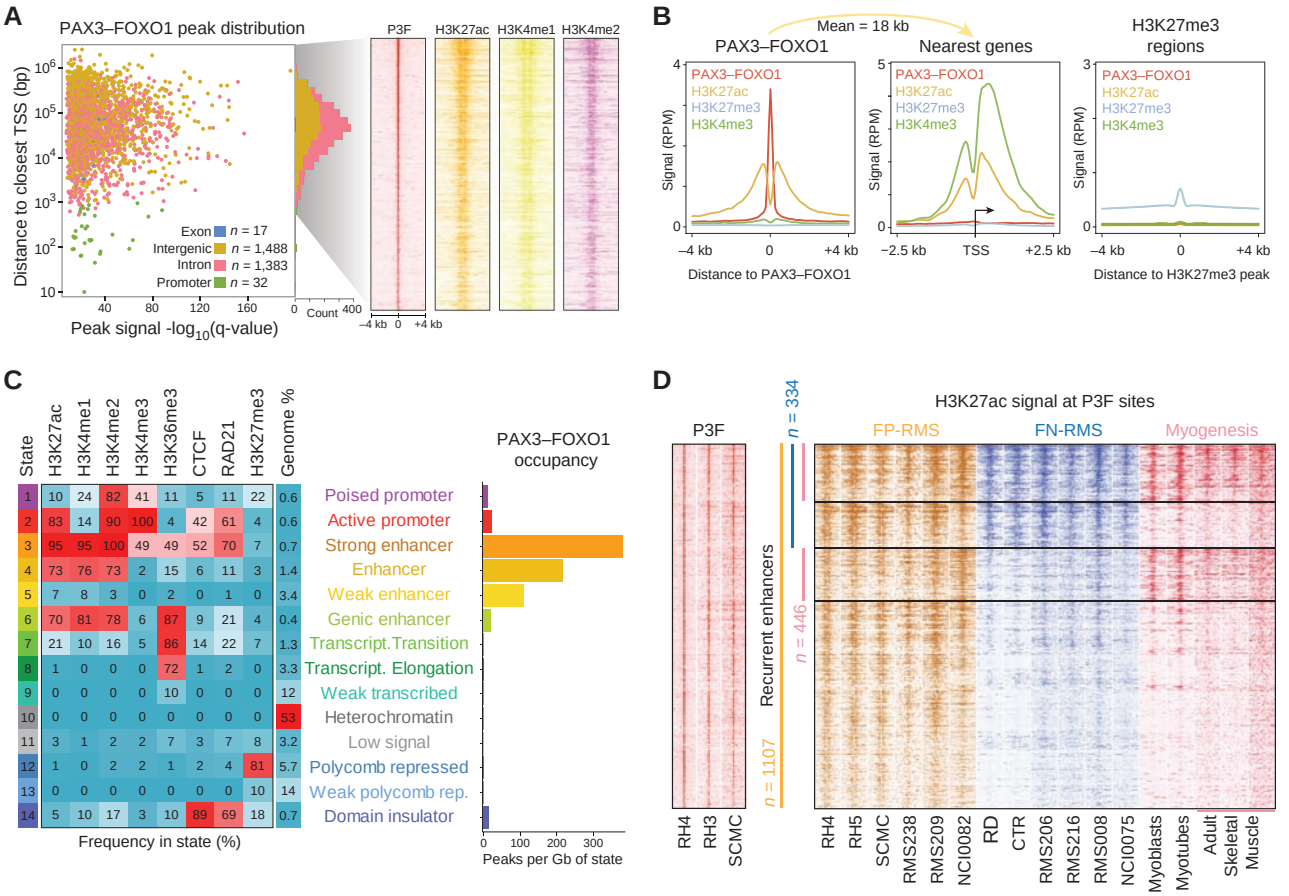

Figure 1. Chromatin state mapping pinpoints PAX3-FOXO1 (P3F) in active enhancers. **A,** PAX3-FOXO1 peak distribution and heat maps of PAX3-FOXO1, H3K27ac, H3K4me1, and H3K4me2 at distal regulatory elements in PAX3-FOXO1-bearing cell line (RH4). The scatter plot is accompanied by a histogram showing the amounts of PAX3-FOXO1 in intronic, exonic, intergenic, or promoter-proximal sites. TSS, transcription start site. Rows are centered around PAX3-FOXO1 peaks and extended 4 kb in each direction, sorted by PAX3-FOXO1 signal strength. **B,** PAX3-FOXO1, H3K27ac, H3K27me3, and H3K4me3 signal at PAX3-FOXO1 peaks (left), genes nearest to PAX3-FOXO1 peaks (center), and Polycomb-repressed chromatin (right). Mean distance of PAX3-FOXO1 to its nearest genes (18 kb) is indicated. **C,** Chromatin states in FP-RMS cells (left) and abundance of PAX3-FOXO1 peaks per Gb of each state (right). States were discovered *de novo* using ChIP-seq data for all histone marks (plus, CTCF and RAD21) with the hidden Markov modeling algorithm chromHMM, which bins the genome into states by recurring patterns. Frequency corresponds to the probability of each mark being present in a given state. **D,** High-confidence PAX3-FOXO1 sites bound to enhancers recurrent ($n = 1,107$) in FP-RMS cell lines and tumors, some of which are shared with FN-RMS ($n = 334$) and/or myogenic cells and tissue ($n = 446$).

RESULTS

PAX3–FOXO1 Establishes Active Chromatin at Distal Enhancers

The hallmark reciprocal translocation of chromosomes 2 and 13 (14) has coding potential for two fusion proteins (Supplementary Fig. S1A), but the only expressed allele has the 5′ end of PAX3 (DNA binding domain) and the 3′ transactivation domain of FOXO1 (Supplementary Fig. S1B). To gain insight into the epigenetic consequences of PAX3–FOXO1, we mapped the landscape of active and repressive histone marks by sequencing DNA enriched by chromatin immunoprecipitation (ChIP-seq) from a patient-derived fusion-positive FP-RMS cell line, RH4. Genome-wide, PAX3–FOXO1 resided predominantly (99%) in sites more than 2.5 kb distal from the nearest transcriptional start site (15), all of which harbored active enhancer marks (Fig. 1A), including acetylation at histone 3 lysine 27 (H3K27ac) and H3 lysine 4 mono/dimethylation (H3K4me1 and H3K4me2), but not the active promoter-associated mark H3K4me3 (Fig. 1B). PAX3–FOXO1 sites showed no evidence of poised and repressed chromatin, as demarcated by Polycomb-deposited trimethylation of H3 lysine 27 (H3K27me3), and, conversely, regions marked by H3K27me3 lacked both PAX3–FOXO1 and H3K27ac (Fig. 1B). Because histone marks are deposited in a combinatorial fashion, we defined reoccurring patterns associated with various chromatin functional states (16), providing the first epigenomic map upon which to overlay PAX3–FOXO1 occupancy (Fig. 1C). We found PAX3–FOXO1 most frequently occupied the strong enhancer chromatin state (Fig. 1C), exemplified by known PAX3–FOXO1 target *FGFR4* (Supplementary Fig. S1C) and oncogenes *MYC, ALK,* and *MET* (Supplementary Fig. S1D).

Published OnlineFirst April 26, 2017; DOI: 10.1158/2159-8290.CD-16-1297

Binding of PAX3–FOXO1 to Enhancers of High Disease and Biological Relevance

To identify which of these PAX3–FOXO1 sites are of high disease relevance, we first performed ChIP-seq in two additional cell FP-RMS lines, RH3 and SCMC, to identify a broader set of targets and recurrent sites. There were 1,783 peaks shared among 2 and 555 peaks for all 3 cell lines, and these showed the most statistical significance and largest signal per peak (Supplementary Fig. S1E–S1F). We next mapped a key histone marker of enhancers H3K27ac in a panel of FP-RMS cell lines and tumor samples, and found 1,107 high-confidence PAX3–FOXO1 sites occupying enhancers in one or more of the cell lines (Fig. 1D; Supplementary Table S1). We then generated a similar map of enhancers in FN-RMS cell lines and tumors, which shared 334 enhancer loci. Four hundred forty-six PAX3–FOXO1-bound enhancers were also present in myogenic samples. The enhancers unique to FP-RMS were enriched, by GREAT ontology (17), in pathways involved in early development, whereas shared enhancers were enriched for late muscle differentiation (Supplementary Fig. S1G).

PAX3–FOXO1 acting across large one-dimensional sequence distances has confounded target gene identification. Previous reports to identify these targets were based on either changes in gene expression (11, 12, 18) or proximity of a gene to a PAX3–FOXO1 peak in a single-cell line with no consideration of expression or chromatin context (15). We therefore identified high-confidence PAX3–FOXO1 target genes by a series of criteria (see Supplementary Fig. S2A–S2B and Supplementary Methods). In brief, we (i) used only PAX3–FOXO1 bound to enhancers that were recurrent in cell lines and tumors (from Fig. 1D), (ii) selected for expressed genes, as PAX3–FOXO1 was found only in active chromatin states, (iii) excluded nearby expressed genes if they were not found within the same topologically associated domain (TAD; predicted by HiC data, ref. 19) as the PAX3–FOXO1 bound enhancer, and (iv) included the maximally expressed gene within each TAD harboring PAX3–FOXO1. Using this approach, we found 1,010 high-confidence targets, 678 of which were novel, and, of note, 439 were significantly reduced by short hairpin RNA (shRNA) knockdown of PAX3–FOXO1 for 48 hours (Supplementary Fig. S2C–S2G). Novel targets included oncogenes ($n = 24$), TFs ($n = 53$), and several imprinted genes ($n = 7$; Supplementary Table S2).

Thus, our data support the hypothesis that PAX3–FOXO1 enables transcription by directing active chromatin marks to distal enhancers surrounding oncogenes, imprinted and myogenic genes.

FP-RMS Tumors and Cell Lines Possess a Myogenic Transcriptional Program

TFs act on enhancers in a stoichiometric manner, and those that are expressed at unusually high concentrations and are able to bind a majority of enhancers are defined as master transcription factors (MTF) and key determinants of cell fate (1). Although PAX3–FOXO1 is the primary driver, as an MTF it is likely to not work alone. We identified other MTFs that were overexpressed compared with normal tissues ($P < 10^{-30}$), had consistently high levels (average FPKM >20), were super enhancer regulated, and had significant motif enrichment in enhancers (Fig. 2A). To do this, we first compared the RNA-sequencing (RNA-seq) profiles of RMS primary tumors

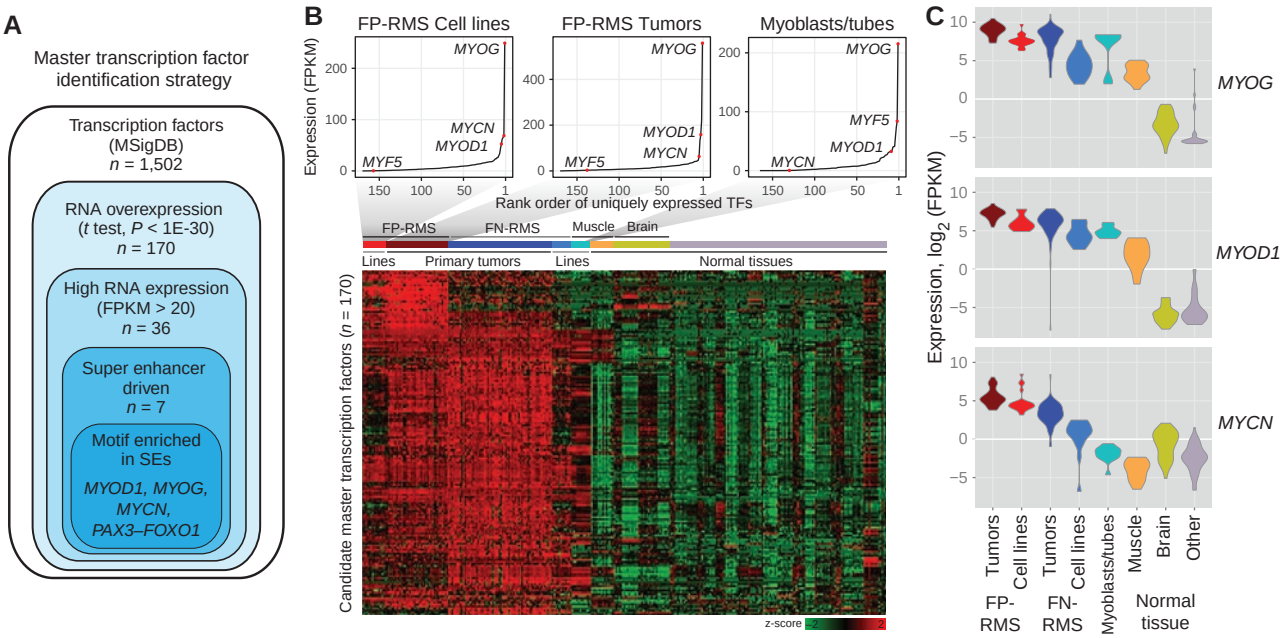

Figure 2. Transcriptional and epigenetic landscapes of RMS cell lines and primary tumors. **A,** Strategy for identifying MTFs from RNA-seq and ChIP-seq datasets. **B,** RNA sequencing of patient RMS tumors reveals a set of candidate MTFs ($n = 170$) identified by filtering for significant overexpression ($P < 1E-30$) across one or both RMS subtypes, and removal of general TFs (those ubiquitously expressed in all tumors and tissues). Candidates were ranked and plotted based on average overall expression in fusion-positive RMS (cell lines and tissues) and myoblasts/myotubes. **C,** Violin plots of gene expression of *MYOD1*, *MYOG*, and *MYCN* across RMS, myoblasts, myotubes, and normal tissues. (*continued on next page*)

Published OnlineFirst April 26, 2017; DOI: 10.1158/2159-8290.CD-16-1297

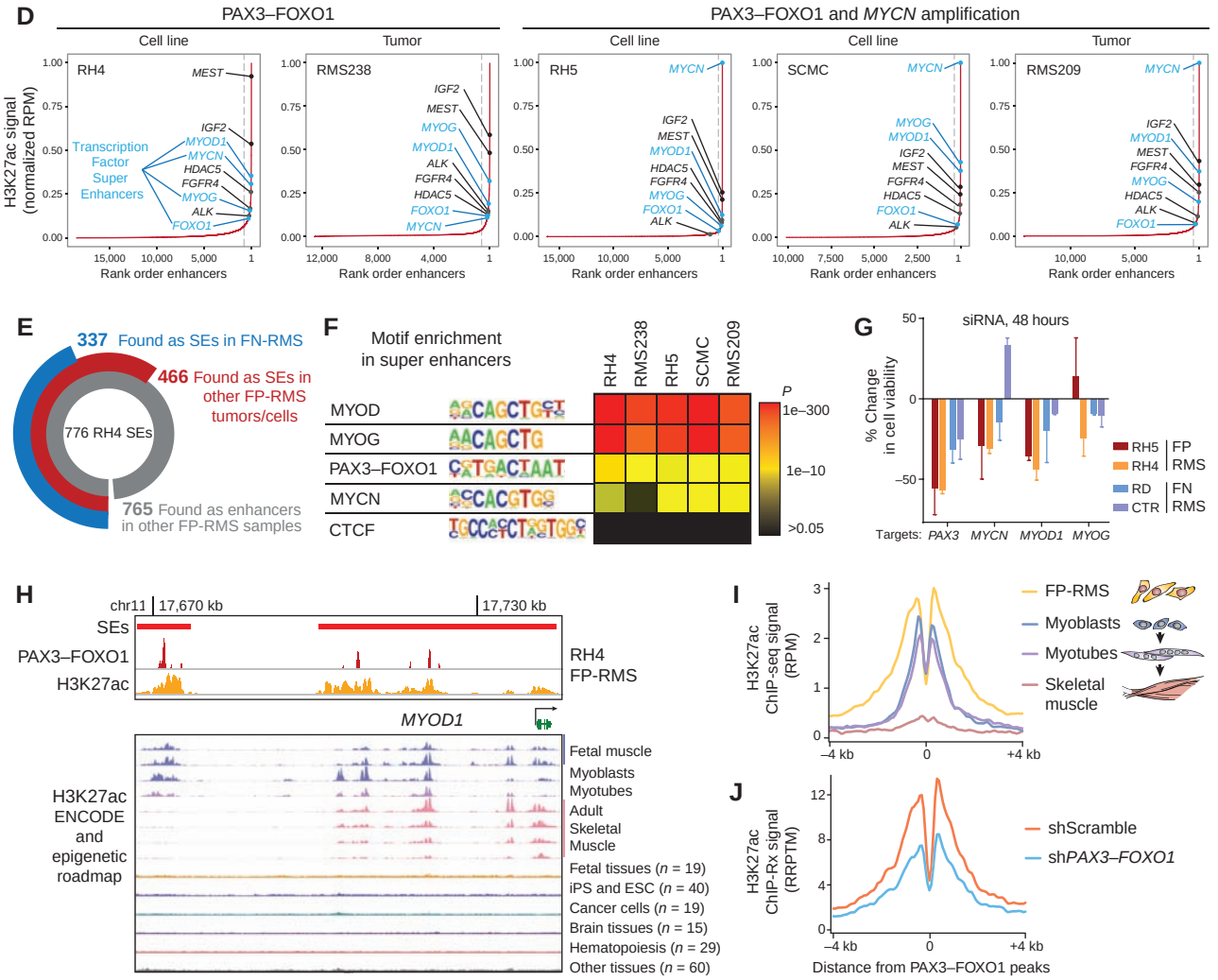

Figure 2. (***Continued***) **D,** H3K27ac binding at distal enhancers ranked by increasing signal in cell lines and primary tumors bearing PAX3–FOXO1 translocation. Super enhancers (SE) were identified as those beyond the inflection point where rapid increase in signal is observed (indicated by a dashed gray line). TF genes associated with SEs are indicated in blue. **E,** Number of RH4 SEs which occur as enhancers ($n = 765$) and SEs in FP-RMS cell lines or tumors ($n = 466$), or which also appear in FN-RMS as SEs ($n = 337$). **F,** Enrichment of known recognition sequences for MYOD, MYOG, PAX3–FOXO1, and MYCN, compared with no enrichment for CTCF in FP-RMS SEs. **G,** Reduction in cell viability upon siRNA against PAX3, MYOD1, MYOG, and MYCN in FN-RMS (RD, CTR) and FP-RMS (RH4, RH5) cells. *PAX3* siRNAs targeted the first few exons, which are intact in the fusion PAX3–FOXO1. CellTiter-Glo was used to measure viability, and all data were normalized to cells treated with scrambled siRNA. Bars show median (error bars = range) of 3 independent siRNA sequences (RH5, RD, and CTR) or 2 independent sequences (RH4). Experiments were performed at 48 hours of transient siRNA transfection. **H**, *MYOD1* enhancers are bound by PAX3–FOXO1 and loaded with active histone mark H3K27ac in RH4 cells (top) and are progressed through myogenesis (middle) and are absent in other cell and tissue types (overlapping plots, bottom). ENCODE and Epigenome Roadmap data tracks are provided at WashU Epigenome browser session http://epigenomegateway.wustl.edu/browser/?genome=hg19&session=IHfj0MDWoA&statusId=728028850. iPS, induced pluripotent stem cells; ESC, embryonic stem cells. **I,** Myogenic enhancers at PAX3–FOXO1 binding sites diminish through muscle differentiation. RPM, reads per million mapped reads. FP-RMS signal is from RH4 cells, myoblasts, and myotubes are from ENCODE; skeletal muscle data are from a normal tissue sample. **J**, Same enhancer locations as **I** interrogated for H3K27ac signal in RH4 cells treated with shRNA for 48 hours (shScramble or sh*PAX3–FOXO1*). Signal was normalized to spike-in *Drosophila* reads (ChIP with reference exogenous genome, ChIP-Rx) and are plotted as reference adjusted reads per 10 million (RRPTM).

($n = 103$) and cell lines ($n = 37$; ref. 3) with normal human organ tissues ($n = 188$). This identified a consistently overexpressed core of 170 TFs, suggesting a convergent underlying epigenetic state (Fig. 2B; Supplementary Table S3). The TFs, of these RMS tumors resembled those of myoblasts and myotubes (20), and all have remarkably high levels of the lineage-determining TFs *MYOD1* and *MYOG* as compared with other normal tissues (Fig. 2B and C). Unlike myoblasts, *MYF5* is typically missing from RMS, although when present it appears to be mutually exclusive with *MYOD1* (21). Another important divergence from normal myogenesis was high expression of the transcriptional amplifier *MYCN* (Fig. 2C), a known target of PAX3–FOXO1 (15). *MYCN* expression was generally higher in FP-RMS tumors whereas *MYC* expression was higher in FN-RMS; however, many tumors had expression of both, such that the sum was consistently high for all

Published OnlineFirst April 26, 2017; DOI: 10.1158/2159-8290.CD-16-1297

patients with RMS (Supplementary Fig. S3A and S3B). RMS cell lines faithfully recapitulated the transcriptional commitment of the primary tumors to these candidate MTFs.

Super Enhancer Analysis in RMS Cell Lines and Primary Tumors Implicates MYOD, MYOG, and MYCN as Master Regulators

A small fraction of active enhancers acquire a large fraction of transcriptional machinery and active chromatin marks (frequently demarcated by H3K27ac), and have been defined as super enhancers (22, 23). These regions are cell type–specific and control expression of cell identity genes in normal tissues and oncogenes in cancer. Although super enhancers may simply be clusters of additive enhancers (24), they nevertheless capture the most active enhancers associated with the core regulatory MTF circuitry (25). In RMS cells bearing PAX3–FOXO1 (RH4), we identified 776 super enhancers (ranked as the top 4% of all enhancers with 38% of the total H3K27ac signal) that were associated with hallmark RMS genes such as *IGF2, FGFR4, ALK, MYOD1, MYOG*, and *MYCN* (Fig. 2D). To evaluate the clinical relevance of our cell lines as models for enhancer architecture, we mapped super enhancers by H3K27ac signal in a set of 3 FP-RMS and 5 FN-RMS primary tumors. We found concordance of super enhancers among PAX3–FOXO1-driven cell lines and tumors (Fig. 2D). *MYCN* had the highest ranked super enhancer in *MYCN*-amplified cell lines and tumors. A common epigenetic landscape was found in FN-RMS cell lines and tumors (Fig. 2E; Supplementary Fig. S3C), which differed from FP-RMS at super enhancers such as *FOXO1, MYCN,* and *ALK* (Supplementary Fig. S3D). *MYCN* itself possessed a remarkable 5 super enhancers within the surrounding TAD structure (HiC data from ref. 19), which circularized chromatin conformation capture followed by sequencing (4C-seq) revealed all physically interact not only with *MYCN*, but also with each other (Supplementary Fig. S3E). Motif analysis at super enhancer sites revealed a highly significant enrichment of MYOD, MYOG, PAX3–FOXO1, and MYCN recognition sequences (Fig. 2F; Supplementary Table S4). Importantly, we observed a consistent reduction in cell viability in 2 FP-RMS cell lines when *PAX3, MYCN,* or *MYOD1* were targeted by siRNAs (Fig. 2G). The suppression was greater in FP-RMS than in FN-RMS, further implicating them as essential master regulators of FP-RMS.

MYOD alone can reprogram cells into myogenesis (26), and it is intriguing that FP-RMS tumors remain undifferentiated despite extremely high levels of the myogenic TFs MYOD and MYOG. Given this, we hypothesized that PAX3–FOXO1 acts as a strong chromatin activator, preventing decommissioning of distal enhancers that are lost during normal muscle differentiation. To study this, we analyzed active enhancer mark H3K27ac (27) around PAX3–FOXO1 peaks, in myoblasts, myotubes, and muscle tissue. We found that enhancers controlling *MYOD1* decreased during muscle differentiation and were absent in other tissues and cell lines, whereas in FP-RMS cells the *MYOD1* enhancer locus was expanded and contained multiple PAX3–FOXO1 peaks (Fig. 2H). Genome-wide, myogenic enhancers found at PAX3–FOXO1 peaks decreased during the transition from myoblasts to myotubes to skeletal muscle (Fig. 2I). Genes associated with these decommissioned enhancers included TFs (*MYCN, MYC, MSC,* and *MYOD1* itself) and epigenetic modulators which have known involvement in regulating self-renewal, embryonic development, muscle development, and chromatin organization (Supplementary Table S5), and imprinted genes involved in mesoderm development (*MEST, IGF2*). These data suggest that PAX3–FOXO1 may induce a myoblastic state by maintaining active chromatin at enhancers controlling these genes. This was further supported by the observation that after 48 hours of PAX3–FOXO1 knockdown these enhancers lost substantial H3K27ac signal as measured by ChIP with reference exogenous genome (ChIP-Rx; ref. 28; Fig. 2J).

PAX3–FOXO1 Collaborates with MYOG, MYOD, and MYCN at Super Enhancers

Our results indicated that PAX3–FOXO1 activates the MTFs MYOD, MYOG, and MYCN, and together with these, establishes the epigenome and transcriptome of FP-RMS (Fig. 3A). To validate this, we performed ChIP-seq on MYOD, MYOG, and MYCN and studied the extent of their collaboration with PAX3–FOXO1 to shape the enhancer landscape. We observed widespread binding of MYOD, MYOG, and MYCN, each covering 9 to 15 times as much genomic space as PAX3–FOXO1 (Supplementary Fig. S4A). When all four MTFs were colocalized in the genome they harbored greater signal of active histone marks, especially H3K27ac (Supplementary Fig. S4B–S4C). Enhancers with all 4 MTFs were frequently super enhancers (Fig. 3B). Almost every super enhancer was occupied by 3 or more of these MTFs, unlike typical enhancers (Fig. 3C). Super enhancers spanned a median of 23.5 kb (compared with 1.2 kb for typical enhancers) and exhibited a higher load of MTFs only when considering their constituent peaks (Fig. 3C; Supplementary Fig. S4D). PAX3–FOXO1 occupied only a small fraction of typical enhancers (5%) but many super enhancers (47%). MYOG, MYOD, and MYCN are bound to almost every super enhancer, and although PAX3–FOXO1 was found in only half of all super enhancers, its preference for super enhancers over typical enhancers was profound (Fig. 3D and E). Super enhancer–associated genes were transcribed at significantly higher levels than typical enhancer genes, and enhancers occupied by 4 MTFs were found to be most highly transcribed compared with enhancers with fewer MTFs (Fig. 3F).

The genome-wide placement of super enhancers was strongly determined by the myogenic MTFs recognizing the underlying DNA sequence CAGCTG (Fig. 2F), but interestingly there was low enrichment of the canonical MYCN motif CACGTG (Supplementary Table S4). More than 80% of MYCN peaks were distal (Supplementary Fig. S4E), and *de novo* motif analysis discovered MYCN prefers a myogenic E-box sequence (RRCAGCTG) nearly identical to that of MYOG and MYOD (Supplementary Fig. S4F). Thus, MYCN may behave in a manner akin to MYC acting as a general transcriptional amplifier, following to locations opened by more sequence-specific (and lineage determining) TFs (29, 30).

Master (or lineage determining) TFs are predicted to maintain cell identity by mutual and self-reinforcement, creating autoregulatory feed-forward loops (1, 31). We found this to be the case in FP-RMS, where the super enhancers controlling *PAX3–FOXO1, MYOD,* and *MYCN* contain all four of these MTFs (Fig. 3G). The *MYOG* super enhancer is bound by all except the PAX fusion, consistent with the logic and timing of normal

Published OnlineFirst April 26, 2017; DOI: 10.1158/2159-8290.CD-16-1297

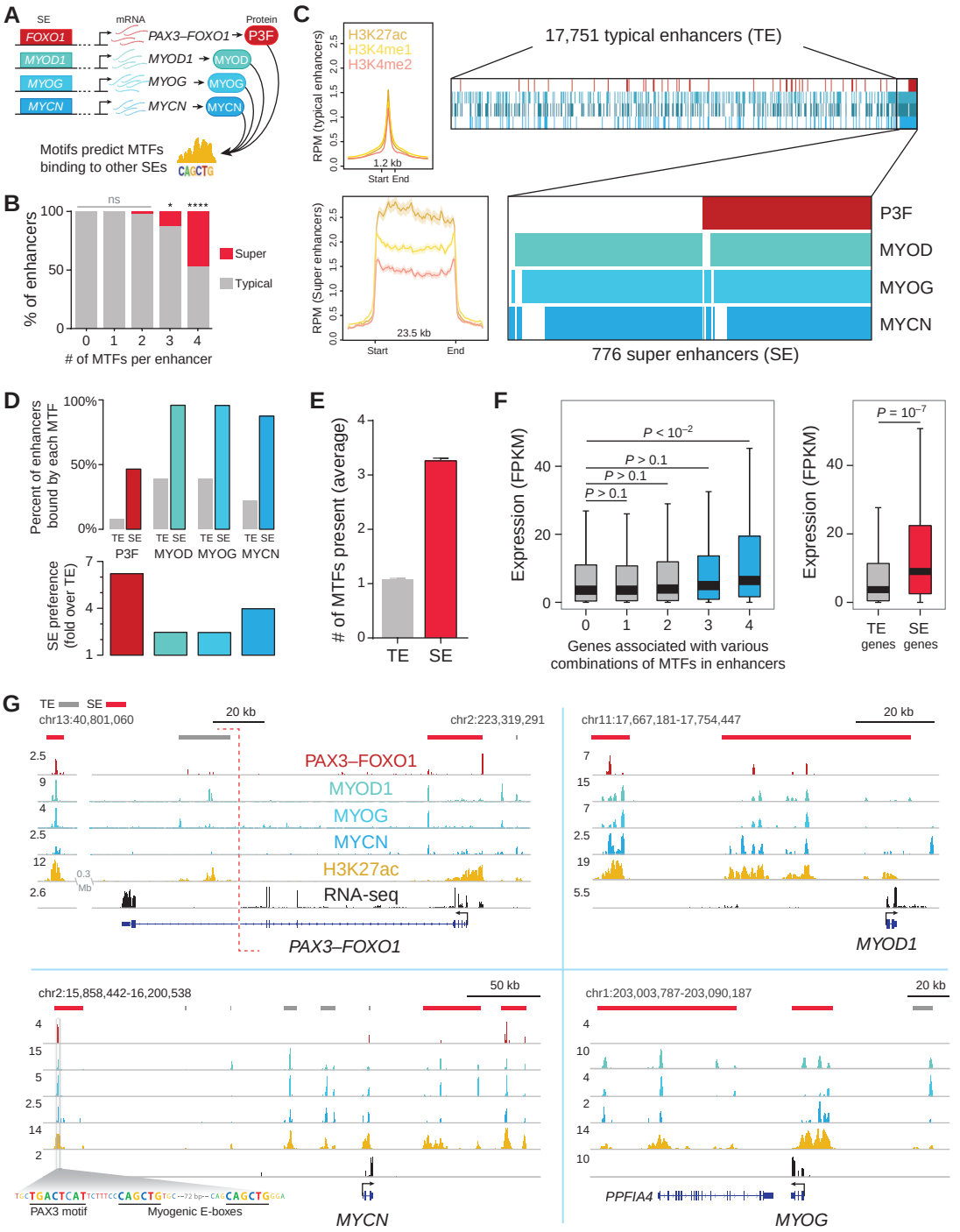

Figure 3. PAX3–FOXO1 (P3F) collaborates with MTFs MYOD, MYOG, and MYCN at super enhancers (SE). **A,** Characteristics enlisting candidate MTFs in FP-RMS: SE-driven, high expression, with motifs enriched across all SEs predicting TF binding. **B,** The percentage of enhancers (divided into groups by the number of MTFs therein) which classify as either typical (TE) or super. Null hypothesis (that the % of super enhancers does not depend on number of MTFs present in an enhancer) was evaluated with Fisher exact test; *, $P < 0.04$; ****, $P < 0.0001$; ns, not significant. **C,** Left, read density profiles of H3K27ac, H3K4me1, and H3K4me2 at regions of TE and SE architecture. Median enhancer length is indicated. Right, collaborative co-occupancy of MTFs in TEs and SEs. Presence of each MTF at enhancer is indicated by the respective colors. **D,** Top, enhancer occupancy of each MTF at TEs or SEs. Bottom, fold enrichment of SEs over TE for each MTFs. **E,** Average number of MTFs per enhancer type. Error bars show 95% confidence interval. **F,** Expression of genes associated with enhancers of various MTF combinations (left) or SE and TE genes (right), associated by proximity. RNA-seq reported as FPKM, fragments per kilobase of transcript per million mapped reads. Error bars, 95% confidence interval. *P* values calculated by Welch unpaired *t* test. **G,** Mutual and self-reinforcement of MTFs via SEs for *PAX3–FOXO1*, *MYOD1*, *MYCN*, and *MYOG*. Tracks show signal in RPM, reads per million mapped reads. TEs are indicated by gray bars and SEs by red. To illustrate an example of multiple adjacent motifs presence within SEs, we have zoomed in on the PAX3 and MYOD/MYOG motifs present upstream of *MYCN*.

Published OnlineFirst April 26, 2017; DOI: 10.1158/2159-8290.CD-16-1297

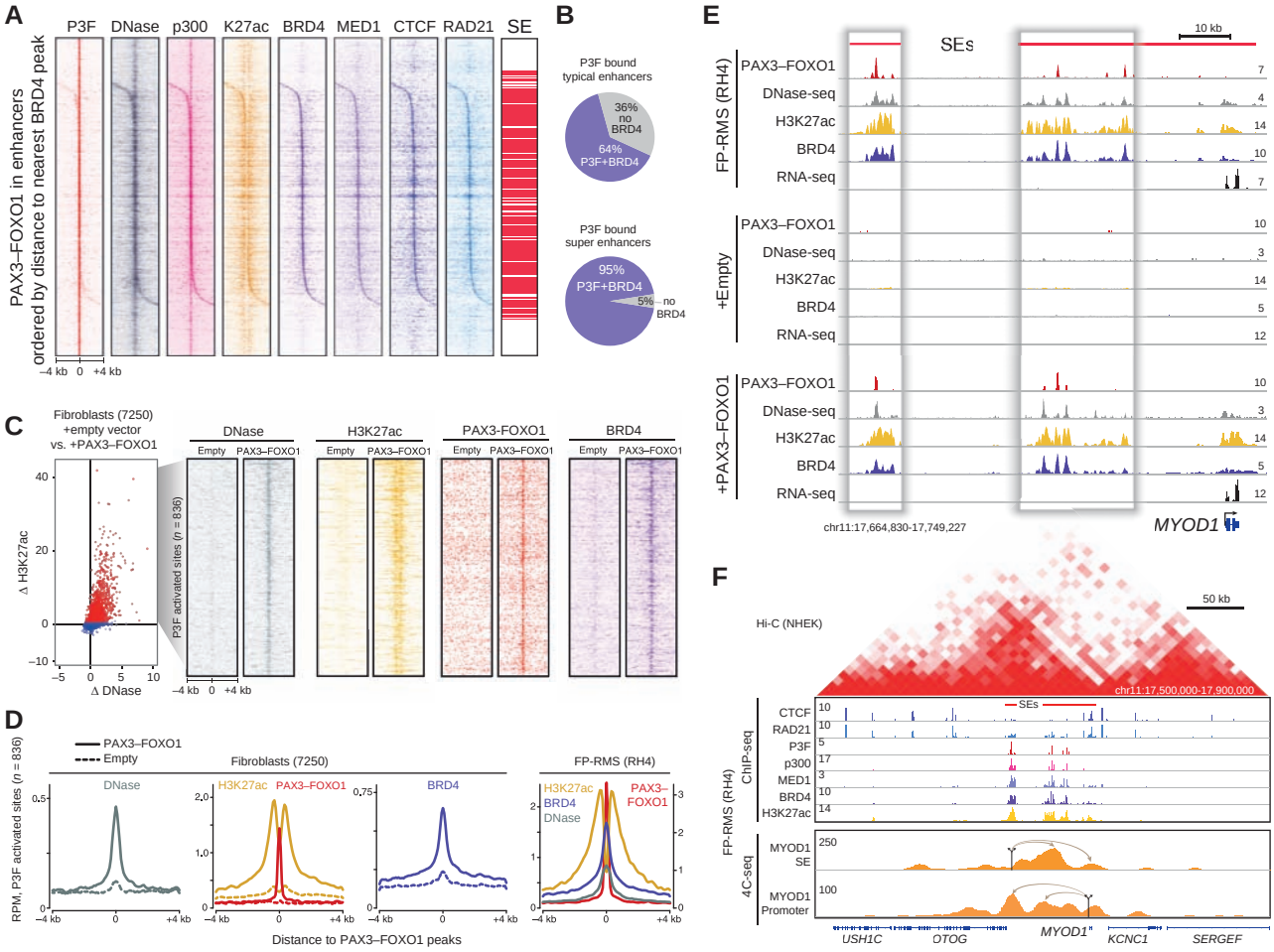

Figure 4. PAX3–FOXO1 (P3F) opens chromatin and recruits BRD4 at looped enhancers. **A,** Heat maps of PAX3–FOXO1, DNase, p300, H3K27ac, BRD4, MED1, CTCF, and RAD21 signal at PAX3–FOXO1 peaks, ranked by distance to closest BRD4 peak in RH4. Rug graph indicates which locations (red lines) are within super enhancers (SE). **B,** PAX3–FOXO1 and BRD4 co-occupancy at enhancers (top, typical; bottom, super). **C,** Introduction of PAX3–FOXO1 into a fibroblast causes increased sensitivity to DNase, deposition of H2K27ac and recruitment of BRD4 at PAX3–FOXO1 sites. **D,** Metagene analysis of DNase, H3K27ac, PAX3–FOXO1, and BRD4 upon PAX3–FOXO1 introduction into fibroblasts (left) compared with RH4 cells (end right). **E,** Opening of chromatin at the *MYOD1* SE by PAX3–FOXO1 in fibroblasts (7,250) with empty vector (middle) or PAX3-FOXO1 (bottom), compared with RH4 cells (top). **F,** Hi-C profile (top) surrounding *MYOD1* locus from NHEK cells (19) with CTCF, RAD21, PAX3–FOXO1, p300, MED1, BRD4, and H3K27ac ChIP-seq data in RH4 cells. 4C-seq (bottom) from *MYOD1* SE looping to adjacent SE region and *MYOD1* promoter, and vice versa, in RH4 cells. Viewpoints are indicated by arrows anchored to their genomic locations.

myogenesis which successively progresses from dominance of PAX3 to MYOD to MYOG (32). To investigate the contribution of interconnection to gene expression, we used shRNA against each factor followed by RNA-seq. PAX3–FOXO1 depletion disrupted both enhancer acetylation and RNA expression at *MYOD* and *MYCN*, but an indirect increase at *MYOG* (Supplementary Fig. S4G). *MYOG* expression was the most profoundly reduced by knockdown of either MYOD1 or MYCN (Supplementary Fig. S4H and S4I). All MTFs were sensitive to MYCN depletion, and *MYCN* expression was sensitive to depletion of any one factor (Supplementary Fig. S4G and S4H).

BRD4, MED1, and p300 Occupy Key PAX3–FOXO1-Established Super Enhancers

Because PAX3–FOXO1 binds between 10Kb and 1Mb away from nearby promoters, it may mediate its transcriptional impact through chromatin factors looping over long distances. The transactivation domain of FOXO1 is known to recruit the coactivator p300 (33), which enzymatically acetylates histones, leading to binding of additional factors, including BRD4 and Mediator (34). We therefore hypothesized that PAX3–FOXO1 recruits p300 and the other cofactors leading to chromatin remodelling. To test this, we performed ChIP-seq of these components and analyzed their co-occupancy at PAX3–FOXO1-bound enhancers. Our results confirmed that these proteins co-occupy enhancers with PAX3–FOXO1 and were sites of open chromatin as determined by DNase hypersensitivity (Fig. 4A). We observed that p300 followed PAX3–FOXO1 at virtually every site, and BRD4 was co-occupant at the majority of these enhancers (72%), especially those with super enhancer architecture (95% of super enhancers; Fig. 4B). PAX3–FOXO1 locations lacking BRD4 showed no evidence of looping machinery MED1, CTCF, and RAD21,

Published OnlineFirst April 26, 2017; DOI: 10.1158/2159-8290.CD-16-1297

whereas sites bound by BRD4 did (Supplementary Fig. S5A). We predicted that these two PAX3–FOXO1 modes (with and without BRD4; Supplementary Fig. S5B–S5C) would have divergent functional consequences, and found that sites with BRD4 had greatly increased expression from associated genes (Supplementary Fig. S5D), and GREAT ontology analysis showed only BRD4-containing peaks were enriched for FP-RMS gene sets (Supplementary Fig. S5E).

It is unknown if PAX3–FOXO1 itself is capable of inducing *de novo* myogenic enhancer formation. We thus stably expressed PAX3–FOXO1 in a human fibroblast cell line (7250) and studied changes in chromatin and the corresponding changes in gene expression. We found that PAX3–FOXO1 opened the chromatin landscape, as evidenced by an increase in DNA hypersensitivity at enhancers compared with control parental cells (Fig. 4C). Furthermore, PAX3–FOXO1-bound enhancers saw an increase in H3K27ac and recruitment of the acetylated lysine reader BRD4 (Fig. 4C). These sites of opening ($n = 836$) are active, PAX3–FOXO1 bound, and BRD4 loaded in FP-RMS cells (Fig. 4D). Many of the enhancers, such as the super enhancers upstream of *MYOD1* (Fig. 4E), *MYOG,* and *FGFR4* (Supplementary Fig. S5F), are faithfully reconstituted with a size and shape similar to those in RMS. Of the 568 high-confidence PAX3–FOXO1 super enhancers, 349 are recapitulated in these fibroblasts reprogrammed by exogenous PAX3–FOXO1 (Supplementary Fig. S5G).

The proximity of the PAX3–FOXO1-directed super enhancers near *MYOD1* and their coappearance with *MYOD1* transcription led us to predict three-dimensional looping to bring these super enhancers to the promoter. We confirmed these *cis* interactions in RH4, using 4C-seq from two viewpoints (one at the most distal PAX3–FOXO1-bound super enhancer and the other at the *MYOD1* promoter; Fig. 4F). These loops enable the physical interaction of the super enhancer–bound proteins MED1, p300, and BRD4 with the promoter of *MYOD1* to facilitate transcription. The 4C interactions were restrained to the TAD predicted by HiC data (19) in other human cells (Fig. 4F). Thus, PAX3–FOXO1 acts as a pioneering factor, opening chromatin, recruiting coactivators, and driving transcription through looped myogenic super enhancers.

Molecular Sensitivities of PAX3–FOXO1 Tumor Cells Are Associated with Super Enhancers

Our results thus far showed that PAX3–FOXO1 establishes super enhancers not only at myogenic genes but also at multiple druggable oncogenic drivers. Because super enhancers are cell-type restricted, we hypothesized that FP-RMS would be selectively vulnerable to inhibition of these super enhancer–driven pathways. Hence, we determined the landscape of molecular sensitivities in PAX3–FOXO1-positive patient-derived cell lines (RH41, RH5) by dose responses at 48 hours for 1,912 compounds. To deprioritize compounds with nonselective cytotoxicity, we also treated fibroblast cell lines (NIH3T3, 7250, and T9195). Our small-molecule library MIPE4 (35) was assembled to have high mechanistic diversity with an emphasis on clinically relevant compounds (Supplementary Fig. S6A–S6B; Supplementary Table S6). Area under the dose response curve (AUC) was used as the measure of potency, as this metric captures both dose dependence and maximum response (see Supplemental Methods and Supplementary Fig. S6C–S6E). PAX3–FOXO1-expressing cells were selectively sensitive to super enhancer–driven RTKs (FGFR4, IGF2/IGF1R, and ALK) and downstream kinases (PI3K, AKT, and mTOR). Furthermore, transcriptional cofactors (HDACs, BRD4) originally identified as super enhancer–associated proteins (23) were also selective for FP-RMS cells (Fig. 5A–B; Supplementary Fig. S6F). Thus, the identification of super enhancer–associated genes highlighted multiple candidate targets for therapy, which may be a useful approach for other cancers. Of note, FP-RMS was sensitive to inhibiton of BRD4 (36), which has recently shown promising results in RMS tumor models (13), although no molecular explanation in connection with PAX3–FOXO1 has previously been made. Thus, with these new data reported herein that BRD4 lies at an important node in the PAX3–FOXO1 circuitry (Supplementary Fig. S6G), we next sought a mechanistic explanation for the sensitivity.

Cells with PAX3–FOXO1 Are Selectively Sensitive to BET Bromodomain Inhibition

Given the sensitivity of FP-RMS cells to chemical BRD inhibition, we sought to determine if this was attributable to PAX3–FOXO1, or if the FN-RMS subtype was also vulnerable. Thus, we tested an expanded panel of RMS cell lines against 5 structurally diverse BET bromodomain inhibitors (BRDi) and 1 pan-bromodomain inhibitor, Bromosporine. Thienodiazepine inhibitors, JQ1 and the clinical analogue OTX015, were the most potent, with dose response consistently in the nanomolar IC_{50} range for FP-RMS cell lines, whereas most often in the micromolar range in fusion-negative (mutant RAS) RMS lines (Fig. 5C; Supplementary Fig. S7A–S7C). BRD4 inhibition dramatically reduced proliferation over time in PAX3–FOXO1-driven cells, whereas mutant RAS-driven RMS cells were relatively unhindered (Fig. 5D; Supplementary Fig. S7D). JQ1 action was mediated by programmed cell death in a dose-dependent manner (Supplementary Fig. S7E). Importantly, although *MYCN* amplification clearly confers special sensitivity to BRD4 inhibition in neuroblastoma (37), FP-RMS cells with (RH5 and SCMC) or without (RH3, RH4, and RH41) *MYCN* amplification were all sensitive (Fig. 5C). PAX3–FOXO1 also conferred 11-fold increased BRDi sensitivity to fibroblasts (Fig. 5E). Patient-derived xenografts grown in culture further confirmed PAX-fusion vulnerability, and FN-RMS resistance, to JQ1 (Supplementary Fig. S7F).

The JQ1 targets BRD2/3/4 (but not BRDT) are expressed in RMS, but not overexpressed compared with normal tissues (Supplementary Fig. S8A). RNAi screening of bromodomain-containing proteins revealed greatest dependence on *EP300, KAT2A, BRD3,* and *BRD4* (Supplementary Fig. S8B). Among the BET family members, *BRD4* was the most sensitive to genetic depletion, which incurred apoptotic events (Supplementary Fig. S8C–S8E).

PAX3–FOXO1 Requires BRD4 for Function and Stability

The sensitivity of PAX3–FOXO1-driven cell lines to BET inhibition and our data showing co-occupancy of BRD4 in all PAX3–FOXO1-bound super enhancers led us to the hypothesis that PAX3–FOXO1 is dependent on BRD4 to mediate transcription of its target genes. This is consistent with the known role of BRD4 in stimulating transcriptional elongation (38) and previous reports that BRD4 inhibition causes rapid

Published OnlineFirst April 26, 2017; DOI: 10.1158/2159-8290.CD-16-1297

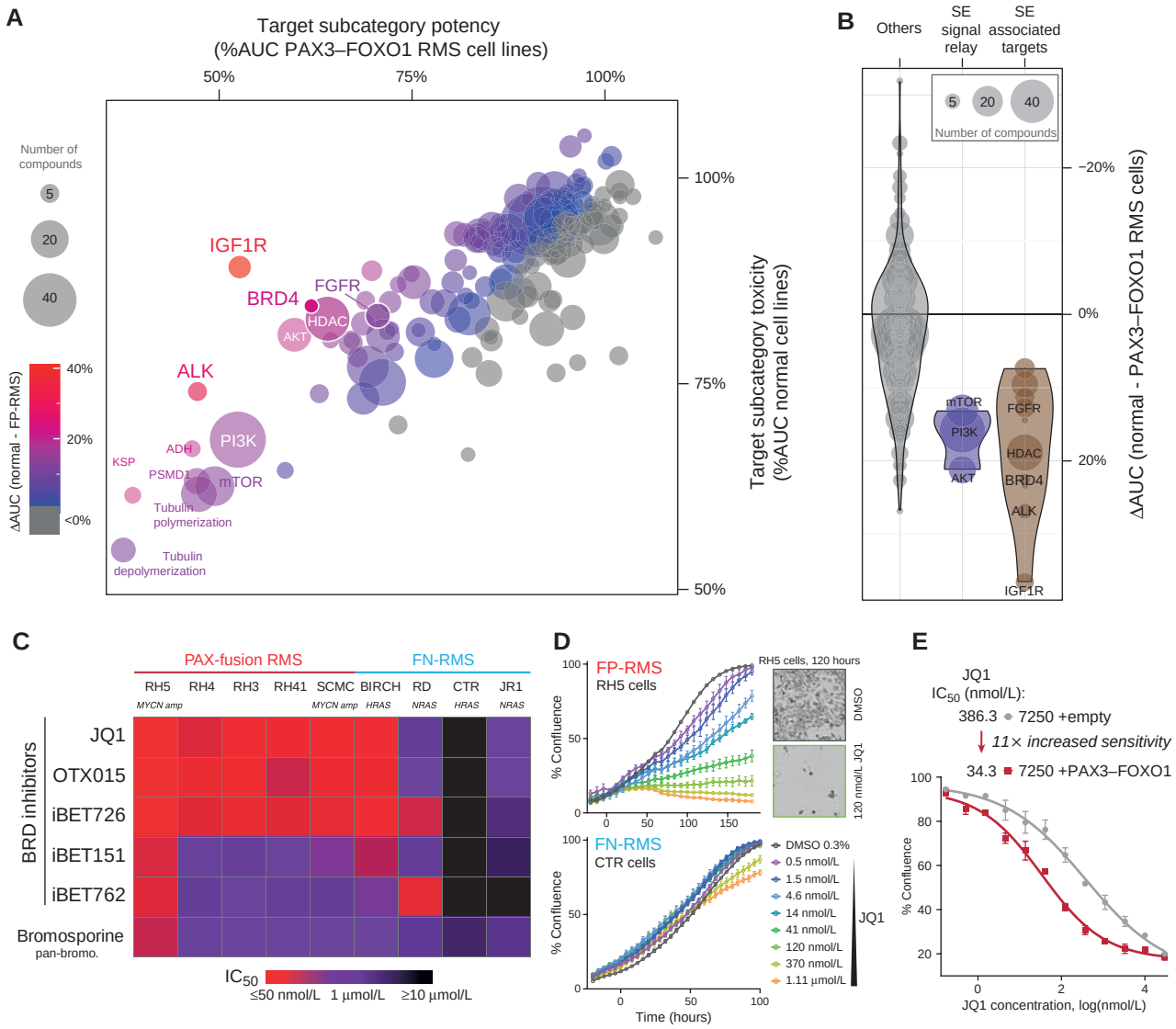

Figure 5. Molecular sensitivity landscape of FP-RMS is enriched in super enhancer (SE)–associated targets including BRD4. **A,** Potency in PAX3–FOXO1 RMS cell lines versus toxicity in normal cell lines measured by dose response and summarized across 240 mechanistically distinct subcategories. The percent area under the dose response curve (%AUC) was averaged for all compounds within a target subcategory. The number of compounds in each category is indicated by the size of the bubble, and the difference in AUC (normal – RMS) is indicated by color scale. **B,** Differential sensitivities against molecules targeting proteins associated with SEs, compared to non-SE targets and SE-signal transduction. Size of the bubble indicates number of molecules against each target. **C,** IC_{50} heat map of 5 BET bromodomain inhibitors and 1 pan-bromodomain inhibitor across 5 PAX-fusion and 4 fusion-negative RMS cell lines. **D,** Growth curves of FP-RMS cells (RH5) and FN-RMS cells (CTR) exposed to increasing concentrations of JQ1 or DMSO. Confluence measured by phase-contrast images every 4 hours over multiple days of treatment. Inset, images of RH5 cells with DMSO or 120 nmol/L JQ1. **E,** PAX3–FOXO1 increases sensitivity of fibroblasts to JQ1.

decommissioning of super enhancers and selective inhibition of super enhancer–driven genes (39). To test this, we compared the fold change in super enhancer–associated versus typical enhancer–associated genes before and after treatment with JQ1 for 6 hours by RNA-seq. Indeed, we found that super enhancer–associated genes in FP-RMS cells were especially sensitive to JQ1, and that this selectivity was also seen upon genetic depletion of PAX3–FOXO1 itself (Fig. 6A; Supplementary Table S7). Gene set enrichment analysis (GSEA) revealed that JQ1 was able to selectively downregulate PAX3–FOXO1 target genes, with enrichment mirroring knockdown of PAX3–FOXO1 (Fig. 6B; Supplementary Fig. S9A–S9B). Many key super enhancer–driven TFs and PAX3–FOXO1 targets were suppressed whereas cell-cycle arrest and apoptosis genes were upregulated by JQ1 (Fig. 6C). The known sensitivity of MYC family proteins to BRD inhibition was seen at both the transcript and protein levels (Fig. 6C; Supplementary Fig. S9C). Coordinately, master regulators MYOD and MYOG were also reduced at the protein level upon JQ1 treatment in a dose- and time-dependent fashion (Fig. 6D; Supplementary Fig. S9D). These effects were not a consequence of reducing PAX3–FOXO1 transcription, as evidenced by exon-level expression in RNA-seq data (Fig. 6E).

Although ChIP-seq evidenced that BRD4 and PAX3–FOXO1 are co-occupant spatially, it was unclear whether they bind

Published OnlineFirst April 26, 2017; DOI: 10.1158/2159-8290.CD-16-1297

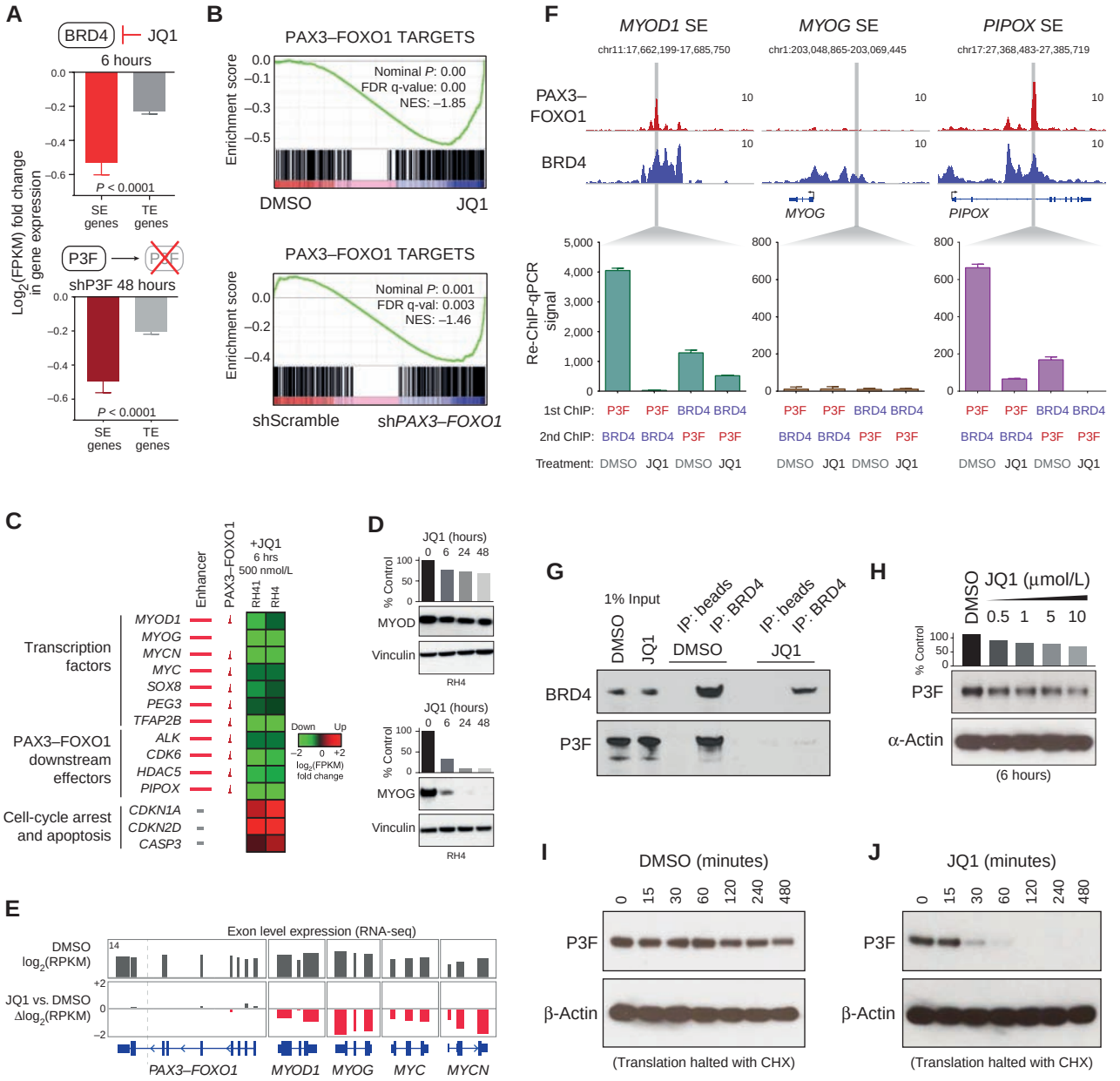

Figure 6. JQ1 selectively ablates PAX3–FOXO1-driven transcription and BRD4 interaction. **A,** Selective disruption of super enhancer (SE) genes upon BET bromodomain inhibition (top) or inducible shRNA depletion of PAX3–FOXO1 (P3F; bottom) in RMS cells (RH4). Fold change in gene expression calculated by comparison with log2 of FPKM in controls (DMSO, scramble shRNA). Error bars show the 95% confidence interval. *P* values calculated by Welch's unpaired *t* test. **B,** GSEA revealed the inhibition of PAX3–FOXO1 fusion gene targets, both by JQ1 and PAX3–FOXO1 knock down. NES, normalized enrichment score; FDR, false discovery rate. Genes used were high-confidence PAX3–FOXO1 targets with recurrent enhancers in 83–100% of FP-RMS samples, as reported in Supplementary Table S2. **C,** mRNA expression alterations of SE (red bar) and PAX3–FOXO1 (red peak) targets after 6 hours of 500 nmol/L JQ1 treatment in RH41 and RH4 cells. Heat map indicates the $\log_2$ fold change in FPKM. **D,** Protein levels of MYOD and MYOG by immunoblotting of RH4 cell lysates after treatment with JQ1(1 μmol/L) over time. **E,** Exon level expression and fold change in RH4 cells upon JQ1 treatment (6 hours, 500 nmol/L), for *PAX3–FOXO1, MYOD1, MYOG, MYC,* and *MYCN*. *PAX3–FOXO1* expression remains intact upon JQ1 treatment, unlike the other key TFs. **F,** BRD4 and PAX3–FOXO1 localization shown via ChIP-seq (top) and re-ChIP-qPCR in the presence and absence of JQ1 (bottom) at the *MYOD* upstream SE, *MYOG* downstream SE and *PIPOX* intronic SE. RH4 cells were treated for 6 hours with DMSO or 1 μmol/L JQ1. **G,** Coimmunoprecipitation of PAX3–FOXO1 and *BRD4* from RH4 cells treated with DMSO or 1 μmol/L JQ1 for 24 hours. **H,** PAX3–FOXO1 immunoblot after 6-hour treatment of DMSO or JQ1 with increasing concentrations. Bar chart (top) quantization of PAX3–FOXO1 normalized to loading controls (β-actin). **I–J,** Stability of PAX3–FOXO1 protein measured by immunoblotting after halting translation with cycloheximide (CHX) in RH4 cells treated with DMSO or JQ1 (1 μmol/L).

Published OnlineFirst April 26, 2017; DOI: 10.1158/2159-8290.CD-16-1297

chromatin cotemporally. To study this, we performed tandem chromatin immunoprecipitations (re-ChIP) for PAX3–FOXO1 followed by BRD4, and vice versa, in RH4 cells treated with either DMSO or JQ1 for 6 hours. Quantitative PCR of re-ChIP DNA at enhancer sites bound by both PAX3–FOXO1 and BRD4 (*MYOD1, PIPOX*) revealed strong enrichment regardless of ChIP order (Fig. 6F), which was almost completely ablated in JQ1-treated cells. This was not observed at BRD4-only enhancers near *MYOG*. To corroborate this, we performed coimmunoprecipitation with BRD4 and PAX3–FOXO1 in the presence of DMSO or JQ1 (Fig. 6G), and found that JQ1 indeed ablated this endogenous interaction. The PAX3–FOXO1 and BRD4 interaction was seen in both directions using exogenous, tagged versions of these proteins (Supplementary Fig. S9E). Given no alteration in *PAX3–FOXO1* mRNA levels, and only modest reduction in protein levels (Fig. 6H), we suspected JQ1 caused destabilization of the PAX3–FOXO1 protein. Remarkably, the half-life of PAX3–FOXO1 was reduced from >8 hours to 28 minutes with JQ1 compared to DMSO (Fig. 6I–J; Supplementary Fig. S9F) in the presence of cycloheximide to inhibit protein translation. This appears to be an on-target effect, as shRNA against BRD4 also caused PAX3–FOXO1 to decrease at the protein level, but not the transcript level (Supplementary Fig. S9G). Thus, PAX3–FOXO1 interacts with BRD4 at enhancers, and treatment with JQ1 leads to loss of this interaction with rapid degradation of PAX3–FOXO1 protein.

JQ1 Selectively Disrupts PAX3–FOXO1-Driven Transcription to Suppress Tumor Growth *In Vivo*

The *in vitro* sensitivity of FP-RMS to BRD4 inhibition by the rapid and the specific inhibition of PAX3–FOXO1 function indicated this may be an effective therapeutic strategy. To test this, we developed an imagable readout to monitor *in vivo* activity of PAX3–FOXO1 super enhancer affected by drugs administered to mice. We engineered RMS cells to express luciferase (and GFP), controlled either by a constitutively active cytomegalovirus (CMV) promoter or a PAX3–FOXO1-dependent super enhancer (cloned from the intronic *ALK* super enhancer; Supplementary Fig. S10A). This *ALK* super enhancer was consistent in FP-RMS cell lines and tumors, while completely absent in FN-RMS (Supplementary Fig. S3D). No activity was seen from the *ALK* super enhancer construct in FN-RMS cells lacking PAX3–FOXO1 (Supplementary Fig. S10B). The *ALK* super enhancer was suppressed by knockdown of PAX3–FOXO1 or by point mutation of the PAX3–FOXO1 binding motif (Supplementary Fig. S10C–S10E). BRD inhibition with JQ1 suppressed only the *ALK* super enhancer– but not the CMV-driven reporter in a dose-dependent manner *in vitro* (Fig. 7A), whereas the general transcription CDK7 inhibitor THZ1 suppressed both. Use of these reporter lines *in vivo* demonstrated a selective inhibition of PAX3–FOXO1 activity with subsequent significant suppression of tumor growth (Fig. 7B and C; Supplementary Fig. S10F). The tumor-suppressive activity of JQ1 was confirmed in a second FP-RMS cell line xenograft, SCMC (Supplementary Fig. S10G–S10H).

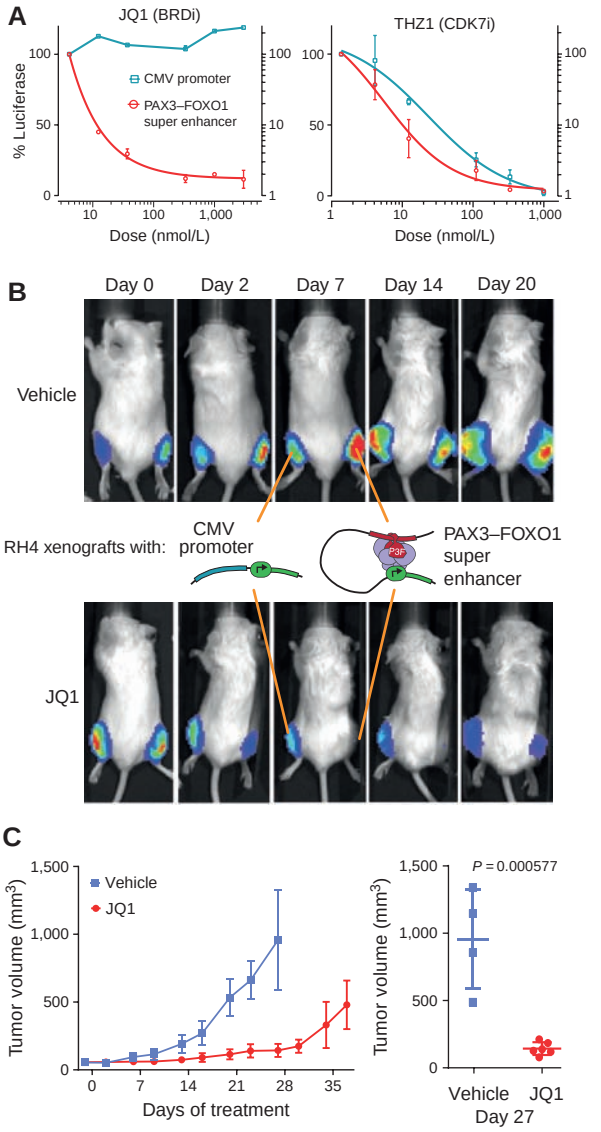

Figure 7. PAX3–FOXO1-dependent super enhancer disruption by BET inhibition *in vivo*. **A,** JQ1 selectively abolishes PAX3–FOXO1-dependent enhancer activity, as measured in PAX3–FOXO1 containing cells (RH4) stably transduced with a lentiviral pGreenFire reporter construct under the control of the PAX3–FOXO1-driven *ALK* super enhancer (SE), while not reducing the CMV-driven expression. CDK7 inhibitor THZ1 inhibits activity of both constructs. PAX3–FOXO1-driven luciferase (red line) is graphed on the left y-axis (linear), and CMV-driven luciferase (blue-green line) is graphed to the right y-axis (log 10 scale). Error bars show standard deviation of duplicate wells, and results are representative of 2 independent experiments. **B,** CMV (left flank) and *ALK* SE (right flank) reporter contructs in RH4 xenografts. JQ1 or vehicle treatment began on day 0 after the first image was taken. **C,** Left, RMS (RH4) tumor growth with vehicle- or JQ1-treated mice. Measurements were taken with caliper and include both CMV and PAX3–FOXO1-SE legs. Right, tumor volume at day 27. *P* value calculated by Welch unpaired *t* test.

DISCUSSION

Pediatric sarcomas that harbor fusion oncogenes are reported with relatively few genomic alterations despite their clinically aggressive behavior (7). Here, we report that the *PAX3–FOXO1* fusion gene directs a profound epigenetic chromatin remodeling in cooperation with the master regulators MYOD, MYOG, and MYCN by the establishment of super enhancers.

By comprehensively charting the first epigenetic landscape of RMS in cell lines and primary tumors, we identified a core

Published OnlineFirst April 26, 2017; DOI: 10.1158/2159-8290.CD-16-1297

commitment to super enhancer regulation of MYOD, MYOG, and MYCN, which in turn drive virtually all densely acetylated enhancer clusters that specify RMS cell identity. Although the concept of a "super" enhancer remains controversial, and may simply be a cluster of additive enhancers (24), we found the classification useful in that the most uniquely and highly expressed candidate MTFs were all driven by super enhancers, often more than one. Our results indicate that PAX3–FOXO1 directly establishes super enhancers to drive itself, *MYOD1*, and *MYCN*, and indirectly establishes a super enhancer (through MYOD and MYCN) to drive *MYOG*. Downstream, these factors are acting in concert, whereas in normal development they act sequentially (32). Virtually all of the discovered super enhancers were bound by 3 or 4 of these MTFs, and associated with the highest levels of gene expression. Remarkably, the presence of MYCN at these hyperactive regions resembles the function of MYC as a general transcriptional amplifier (29) and may underlie transcriptional addiction in FP-RMS (40). Together, these four MTFs cause a profound epigenetic reprogramming, freezing the cells in a myoblastic state with PAX3–FOXO1 as the conductor.

Cell type–specific distribution of BRD4 is accomplished by TF recruitment to their enhancers (41). Our study showed that PAX3–FOXO1 can direct *de novo* recruitment of BRD4 to specific chromatin sites when introduced into human fibroblasts, accompanied by opening of chromatin and acetylation of H3K27 at enhancers proximal to key RMS genes. This purely activating function of PAX3–FOXO1 in pediatric RMS is in sharp contrast to the dual functionality of EWS–FLI1 in pediatric Ewing sarcoma (5), which has both activating and repressive functions depending on genomic context.

The extent to which epigenetic profiling can aid prediction and interpretation of chemical sensitivities remains unclear. Using mechanistically informed drug screening, we observed that molecules targeted to super enhancer–bound coactivators (BRD4 and HDAC; ref. 22) and super enhancer targets (such as the receptor tyrosine kinases FGFR4 and ALK) are the most selectively potent in these FP-RMS cells. These data reinforce previous strategies that showed FP-RMS sensitivity to FGFR and IGF1R inhibition (42, 43) and HDAC inhibitors (4, 44, 45), and add previously unknown mechanistic insights to recently discovered BRD4 vulnerability (13). The observation that super enhancer–associated targets represent key vulnerabilities may be broadly applicable to cancers driven by epigenetic reprograming through super enhancer networks and may be critical to prioritizing combination strategies as well.

The same translocation which causes FP-RMS also creates an Achilles' heel by addicting cells to BRD4. Our data add to the pleiotropic utility of BRD4 inhibitors, with a recent wave of studies in a diverse group of cancers (37, 46–50). The transcriptional impact of BRD4 inhibition appears to be context dependent, where its antitumor effect is linked to dampening one or more master regulators, often including *MYC* in hematologic malignancies (51, 52), *MYCN* in neuroblastoma (37), *POU2AF1* and *PAX5* in DLBCL (47), or *FOSL1* in adenocarcinoma (53), among others. JQ1 was shown recently to be effective in reducing RMS tumor growth by antiangiogenic properties, but the underlying chromatin-based mechanisms or FP-RMS subtype selectivity were not explored (13). Here, we find BRDi caused selective downregulation of *MYC, MYCN, MYOD1, MYOG*, and many other downstream PAX3–FOXO1- and super enhancer–driven genes. Surprisingly, JQ1 did not decrease expression of *PAX3–FOXO1* mRNA at 6 hours, yet it rapidly decreased PAX3–FOXO1 protein stability. This may result from the disruption of interaction between BRD4 and PAX3–FOXO1, discovered by both coimmunoprecipitation and re-ChIP experiments, which JQ1 abrogated within hours of drug exposure. Thus, BRD inhibition by JQ1 leads to significant tumor suppression *in vitro* and *in vivo,* ablating the transcription-driving function of the fusion gene. Indeed, this mechanism may partially explain the antitumor effects seen with BI-2536 in FP-RMS mouse models (54), as this PLK1 inhibitor was recently found to possess a nanomolar inhibition of BET bromodomains (55). Excitingly, this provides a means of selectively drugging PAX3–FOXO1, a long-standing goal of FP-RMS research (56), and provides a new precision therapeutic for treatment of the aggressive PAX fusion–bearing RMS.

METHODS

Cell Lines and Primary Tumors

All cell lines were routinely tested for *Mycoplasma* within one or two passages of each experiment herein, and cell line identities have been ensured by RNA-seq and genotyping. RH4, RH3, RH5, and RH41 were kind gifts from Dr. Peter Houghton (obtained between 2008 and 2010), SCMC from Dr. Janet Shipley (obtained between 2013 and 2015), and RD, CTR, and Birch from Dr. Lee Helman (obtained between 2008 and 2010). CRL7250 and NIH3T3 were obtained from the ATCC (obtained between 2008 and 2010). Validation was performed by DNA fingerprinting AmpFlSTR Identifiler PCR Amplification Kit (Catalog Number 4322288) by Life Technologies. Cell lines were grown in DMEM, 10% FBS, and supplemented with penicillin/streptomycin. Primary RMS cultures established from patient-derived tumor xenografts were collected at the St. Jude Children's Research Hospital (57). Cells were maintained in Neurobasal-A medium (ThermoFisher Scientific) supplemented with 2xB-27 supplement (ThermoFisher Scientific), 20 ng/mL bFGF, and 20 ng/mL EGF (both from PeproTech), and cultured on gelatin-coated plates in 5% CO_2 at 37°C. Primary tumors were acquired via the NCI-coordinated ClinOmics protocol as previously described (58).

ChIP-seq and RNA-seq

ChIP-seq (59) and RNA-seq (3) were performed as previously described. Raw sequencing data and processed files have been made available through Gene Expression Omnibus (GEO) SuperSeries accession number GSE83728, which is comprised of SubSeries accession numbers GSE83724, GSE83725, GSE83726, and GSE83727. Details, including Illumina sequencing and bioinformatic methods, are available in Supplementary Methods.

DNase-seq with 10,000 Cells

Sites of DNase-sensitive chromatin were captured from 10,000 cells as recently described (60) with slight modifications. Briefly, freshly trypsinized cells were resuspended in DMEM, counted in duplicate (Nexcelom Automated Cell Counter), pelleted and resuspended in lysis buffer to achieve 120 µL of 60K cells, which was then divided into 6 replicates (10K cells per tube). DNaseI (Roche 04-716-728-001) was added to the cells (0.25–0.5 units) and incubated for 5 minutes at 37°C. The digestion was halted with 50 µL of stop buffer (9.5 mL H_2O + 100 µL 1M TrisHCl pH 7.4 + 20 µL 5 mol/L NaCl + 200 µL 0.5 mol/L EDTA, with 150 µL 10% SDS and 125 µL proteinase K added just before use). Proteinase K activation at 55°C for 1 hour was followed by DNA purification by column (MiniElute PCR

Published OnlineFirst April 26, 2017; DOI: 10.1158/2159-8290.CD-16-1297

purification kit, Qiagen). Library preparation was performed as with ChIP-seq samples, except that paired-end was used rather than single-end sequencing on the NextSeq 500 (Illumina).

Small-Molecule Compounds

All molecules were dissolved in DMSO to a final concentration of 10 mmol/L and diluted to a final DMSO concentration of <0.03% by volume in DMEM for cell culture experiments. JQ1 was a gift from Jay Bradner (Novartis) and Jun Qi (Dana-Farber Cancer Institute). Bromosporine was provided by Peter Brown of the Structural Genomics Consortium. THZ1 was supplied by Nat Gray (Dana-Farber Cancer Institute). Other bromodomain inhibitors (OTX015, I-Bet-151, I-Bet-762, and I-Bet-726) were generously supplied by the Developmental Therapeutics Program (NCI, NIH).

Time Course of Dose-Response Cell Growth Assay

Dose responses were performed by quantifying percent cell confluence from phase contrast images taken every 4 hours using the IncuCyte ZOOM in 384-well plate format. Dose response was achieved using a range of 12 concentrations from 30 μmol/L to 0.17 nmol/L (dilutions divided by 3) and were performed in triplicate. Cells were plated to achieve 15% confluence at time of drug dosing, and monitored until control (DMSO) wells reached >95% confluence. IC_{50} values were calculated for each time point using the R statistical package drc (https://cran.r-project.org/web/packages/drc/drc.pdf).

Luciferase-Expressing Cells

pGreenFire vector from Systems Biosciences was modified by insertion of a *cis*-regulatory element surrounding the PAX3–FOXO1 binding site within the super enhancer found within the intronic region of the ALK gene (chr2:29880537–29880842). Cloning was performed using pCR2.1–TOPO system from PCR amplification product from genomic DNA of RH4 FP-RMS cells, and shuttled into the pGF1 vector at the EcoRI restriction site upstream of a minimal CMV promoter, which was completely inactive on its own in RMS cell lines. Viral particles were produced in HEK293T cells, harvested, filtered, and pelleted. pGF1 cloning vectors developed at SBI are self-inactivating as a result of a deletion in the U3 region of 3′-LTR. Upon integration into the genome, the 5′ LTR promoter is inactivated, which prevents formation of replication-competent viral particles. Pooled cells were selected using puromycin.

Animal Studies

Animal studies were approved by the National Cancer Institute's Animal Care and Use Committee, and all animal care was in accordance with institutional guidelines. Complete details are reported in Supplementary Methods.

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

Authors' Contributions

Conception and design: B.E. Gryder, M.E. Yohe, H.-C. Chou, J.S. Wei, C.J. Thomas, J. Khan
Development of methodology: B.E. Gryder, H.-C. Chou, Y. Song, R. Rota, X. Wen, S. Sindiri, M. Ferrer, J.F. Shern, C.J. Thomas, J. Khan
Acquisition of data (provided animals, acquired and managed patients, provided facilities, etc.): B.E. Gryder, M.E. Yohe, X. Zhang, J. Marques, M. Wachtel, N. Sen, Y. Song, R. Rota, X. Wen, J.S. Wei, F.G. Barr, S. Das, T. Andresson, M. Lal-Nag, M. Ferrer, C.J. Thomas, J. Khan
Analysis and interpretation of data (e.g., statistical analysis, biostatistics, computational analysis): B.E. Gryder, H.-C. Chou, J. Marques, M. Wachtel, B. Schaefer, N. Sen, A. Cleveland, X. Wen, S. Sindiri, J.S. Wei, F.G. Barr, R. Guha, M. Lal-Nag, M. Ferrer, J.F. Shern, K. Zhao, C.J. Thomas, J. Khan
Writing, review, and/or revision of the manuscript: B.E. Gryder, M.E. Yohe, H.-C. Chou, B. Schaefer, N. Sen, R. Rota, J.S. Wei, F.G. Barr, R. Guha, M. Ferrer, J.F. Shern, K. Zhao, C.J. Thomas, J. Khan
Administrative, technical, or material support (i.e., reporting or organizing data, constructing databases): B.E. Gryder, H.-C. Chou, X. Wen, J. Khan
Study supervision: R. Rota, C.J. Thomas, J. Khan
Other (writing original manuscript): B.E. Gryder
Other (wet-lab investigation, providing materials): A. Gualtieri, S. Pomella

Acknowledgments

The authors are very grateful to C. Thiele, B. Stanton, J. Waterfall, L. Cao, T. Misteli, and C. Vakoc for helpful discussions and review of the manuscript. We are thankful to J. Qi, J. Bradner, N. Gray, P. Brown, the Structural Genomics Consortium, and the Developmental Therapeutics Program (NCI, NIH) for providing small molecules. We also appreciate the help of S. Li, N. Shivaprasad, S. Michael, and M. Skarzynski in providing reagents, cell lines, and antibodies. We thank C. Lin and R. Patidar for sharing scripts and bioinformatic advice.

Grant Support

This project has been funded with federal funds from the National Cancer Institute, NIH, contract no. HHSN261200800001E. A. Gualtieri is a recipient of a Fondazione Veronesi fellowship. R. Rota has been supported by Associazione Italiana Ricerca sul Cancro (AIRC) and PE-2013-02355271.

The content of this publication does not necessarily reflect the views or policies of the Department of Health and Human Services, nor does mention of trade names, commercial products, or organizations imply endorsement by the U.S. government. Frederick National Laboratory is accredited by AAALAC International and follows the Public Health Service Policy for the Care and Use of Laboratory Animals. Animal care was provided in accordance with the procedures outlined in the "Guide for Care and Use of Laboratory Animals" (National Research Council, 2011; National Academies Press, Washington, DC).

Received November 19, 2016; revised March 20, 2017; accepted April 21, 2017; published OnlineFirst April 26, 2017.

REFERENCES

1. Lee Tong I, Young Richard A. Transcriptional regulation and its misregulation in disease. Cell 2013;152:1237–51.
2. Rabbitts T. Chromosomal translocations in human cancer. Nature 1994;372:143–9.
3. Shern JF, Chen L, Chmielecki J, Wei JS, Patidar R, Rosenberg M, et al. Comprehensive genomic analysis of rhabdomyosarcoma reveals a landscape of alterations affecting a common genetic axis in fusion-positive and fusion-negative tumors. Cancer Discov 2014;4:216–31.
4. Chen X, Stewart E, Shelat Anang A, Qu C, Bahrami A, Hatley M, et al. Targeting Oxidative Stress in Embryonal Rhabdomyosarcoma. Cancer Cell 2013;24:710–24.
5. Riggi N, Knoechel B, Gillespie Shawn M, Rheinbay E, Boulay G, Suvà Mario L, et al. EWS-FLI1 utilizes divergent chromatin remodeling mechanisms to directly activate or repress enhancer elements in ewing sarcoma. Cancer Cell 2014;26:668–81.
6. Kadoch C, Crabtree Gerald R. Reversible Disruption of mSWI/SNF (BAF) Complexes by the SS18-SSX oncogenic fusion in synovial sarcoma. Cell 2013;153:71–85.
7. Pizzo PA, Poplack DG. Principles and practice of pediatric oncology. Philadelphia: Lippincott Williams & Wilkins; 2015.

Published OnlineFirst April 26, 2017; DOI: 10.1158/2159-8290.CD-16-1297

8. Missiaglia E, Williamson D, Chisholm J, Wirapati P, Pierron G, Petel F, et al. PAX3/FOXO1 fusion gene status is the key prognostic molecular marker in rhabdomyosarcoma and significantly improves current risk stratification. J Clin Oncol 2012;30:1670–7.
9. Khan J, Simon R, Bittner M, Chen Y, Leighton SB, Pohida T, et al. Gene expression profiling of alveolar rhabdomyosarcoma with cDNA microarrays. Cancer Res 1998;58:5009–13.
10. Davis RJ, Barr FG. Fusion genes resulting from alternative chromosomal translocations are overexpressed by gene-specific mechanisms in alveolar rhabdomyosarcoma. Proc Nat Acad Sci 1997;94:8047–51.
11. Davicioni E, Graf Finckenstein F, Shahbazian V, Buckley JD, Triche TJ, Anderson MJ. Identification of a PAX-FKHR gene expression signature that defines molecular classes and determines the prognosis of alveolar rhabdomyosarcomas. Cancer Res 2006;66:6936–46.
12. Khan J, Bittner ML, Saal LH, Teichmann U, Azorsa DO, Gooden GC, et al. cDNA microarrays detect activation of a myogenic transcription program by the PAX3–FKHR fusion oncogene. Proc Natl Acad Sci U S A 1999;96:13264–9.
13. Bid HK, Phelps DA, Xiao L, Guttridge DC, Lin J, London C, et al. The bromodomain BET inhibitor JQ1 suppresses tumor angiogenesis in models of childhood sarcoma. Mol Cancer Ther 2016;15:1018–28.
14. Barr FG, Galili N, Holick J, Biegel JA, Rovera G, Emanuel BS. Rearrangement of the PAX3 paired box gene in the paediatric solid tumour alveolar rhabdomyosarcoma. Nat Genet 1993;3:113–7.
15. Cao LA, Yu YK, Bilke S, Walker RL, Mayeenuddin LH, Azorsa DO, et al. Genome-Wide Identification of PAX3–FKHR binding sites in rhabdomyosarcoma reveals candidate target genes important for development and cancer. Cancer Res 2010;70:6497–508.
16. Roadmap Epigenomics C, Kundaje A, Meuleman W, Ernst J, Bilenky M, Yen A, et al. Integrative analysis of 111 reference human epigenomes. Nature 2015;518:317–30.
17. McLean CY, Bristor D, Hiller M, Clarke SL, Schaar BT, Lowe CB, et al. GREAT improves functional interpretation of cis-regulatory regions. Nat Biotechnol 2010;28:495–501.
18. Begum S, Emani N, Cheung A, Wilkins O, Der S, Hamel PA. Cell-type-specific regulation of distinct sets of gene targets by Pax3 and Pax3//FKHR 2005;24:1860–72.
19. Rao Suhas SP, Huntley Miriam H, Durand Neva C, Stamenova Elena K, Bochkov Ivan D, Robinson James T, et al. A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. Cell 2014;159:1665–80.
20. Trapnell C, Cacchiarelli D, Grimsby J, Pokharel P, Li S, Morse M, et al. The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. Nat Biotech 2014;32:381–6.
21. Tenente IM, Hayes MN, Ignatius MS, McCarthy K, Yohe M, Sindiri S, et al. Myogenic regulatory transcription factors regulate growth in rhabdomyosarcoma. eLife 2017;6:e19214.
22. Hnisz D, Abraham Brian J, Lee Tong I, Lau A, Saint-André V, Sigova Alla A, et al. Super-enhancers in the control of cell identity and disease. Cell 2013;155:934–47.
23. Whyte WA, Orlando DA, Hnisz D, Abraham BJ, Lin CY, Kagey MH, et al. Master transcription factors and mediator establish super-enhancers at key cell identity genes. Cell 2013;153:307–19.
24. Dukler N, Gulko B, Huang Y-F, Siepel A. Is a super-enhancer greater than the sum of its parts? Nat Genet 2017;49:2–3.
25. Lin CY, Erkek S, Tong Y, Yin L, Federation AJ, Zapatka M, et al. Active medulloblastoma enhancers reveal subgroup-specific cellular origins. Nature 2016;530:57–62.
26. Davis RL, Weintraub H, Lassar AB. Expression of a single transfected cDNA converts fibroblasts to myoblasts. Cell 1987;51:987–1000.
27. Ernst J, Kheradpour P, Mikkelsen TS, Shoresh N, Ward LD, Epstein CB, et al. Mapping and analysis of chromatin state dynamics in nine human cell types. Nature 2011;473:43–9.
28. Orlando David A, Chen Mei W, Brown Victoria E, Solanki S, Choi Yoon J, Olson Eric R, et al. Quantitative ChIP-Seq normalization reveals global modulation of the epigenome. Cell Reports 2014;9:1163–70.
29. Lin Charles Y, Lovén J, Rahl Peter B, Paranal Ronald M, Burge Christopher B, Bradner James E, et al. Transcriptional amplification in tumor cells with elevated c-myc. Cell 2012;151:56–67.
30. Nie Z, Hu G, Wei G, Cui K, Yamane A, Resch W, et al. c-Myc is a universal amplifier of expressed genes in lymphocytes and embryonic stem cells. Cell 2012;151:68–79.
31. Saint-André V, Federation AJ, Lin CY, Abraham BJ, Reddy J, Lee TI, et al. Models of human core transcriptional regulatory circuitries. Genome Res 2016;26:385–96.
32. Hettmer S, Wagers AJ. Muscling in: Uncovering the origins of rhabdomyosarcoma. Nat Med 2010;16:171–3.
33. Kitamura T, Feng Y, Kitamura YI, Chua SC, Xu AW, Barsh GS, et al. Forkhead protein FoxO1 mediates Agrp-dependent effects of leptin on food intake. Nature medicine 2006;12:534–40.
34. Shi J, Vakoc Christopher R. The mechanisms behind the Therapeutic Activity of BET Bromodomain Inhibition. Mol Cell 2014;54:728–36.
35. Jones LH, Bunnage ME. Applications of chemogenomic library screening in drug discovery. Nat Rev Drug Discov 2017;16:285–96.
36. Filippakopoulos P, Qi J, Picaud S, Shen Y, Smith WB, Fedorov O, et al. Selective inhibition of BET bromodomains. Nature 2010;468:1067–73.
37. Puissant A, Frumm SM, Alexe G, Bassil CF, Qi J, Chanthery YH, et al. Targeting MYCN in neuroblastoma by BET bromodomain inhibition. Cancer Discovy 2013;3:308–23.
38. Jang MK, Mochizuki K, Zhou M, Jeong H-S, Brady JN, Ozato K. The bromodomain protein Brd4 Is a positive regulatory component of P-TEFb and stimulates RNA polymerase II-dependent transcription. Mol Cell 2005;19:523–34.
39. Lovén J, Hoke Heather A, Lin Charles Y, Lau A, Orlando David A, Vakoc Christopher R, et al. Selective inhibition of tumor oncogenes by disruption of super-enhancers. Cell 2013;153:320–34.
40. Bradner JE, Hnisz D, Young RA. Transcriptional addiction in cancer. Cell 2017;168:629–43.
41. Roe J-S, Mercan F, Rivera K, Pappin Darryl J, Vakoc Christopher R. BET bromodomain inhibition suppresses the function of hematopoietic transcription factors in acute myeloid leukemia. Mol Cell 2015;58:1028–39.
42. Crose LE, Etheridge KT, Chen C, Belyea B, Talbot LJ, Bentley RC, et al. FGFR4 blockade exerts distinct antitumorigenic effects in human embryonal versus alveolar rhabdomyosarcoma. Clin Cancer Res 2012;18:3780–90.
43. Kim SY, Toretsky JA, Scher D, Helman LJ. The role of IGF-1R in pediatric malignancies. Oncologist 2009;14:83–91.
44. Abraham J, Nuñez-Álvarez Y, Hettmer S, Carrió E, Chen H-IH, Nishijo K, et al. Lineage of origin in rhabdomyosarcoma informs pharmacological response. Genes Deve 2014;28:1578–91.
45. Kutko MC, Glick RD, Butler LM, Coffey DC, Rifkind RA, Marks PA, et al. Histone deacetylase inhibitors induce growth suppression and cell death in human rhabdomyosarcoma in vitro. Clin Cancer Res 2003;9:5749–55.
46. Asangani IA, Dommeti VL, Wang X, Malik R, Cieslik M, Yang R, et al. Therapeutic targeting of BET bromodomain proteins in castration-resistant prostate cancer. Nature 2014;510:278–82.
47. Chapuy B, McKeown Michael R, Lin Charles Y, Monti S, Roemer Margaretha GM, Qi J, et al. Discovery and characterization of super-enhancer-associated dependencies in diffuse large B cell lymphoma. Cancer cell 2013;24:777–90.
48. Dawson MA, Prinjha RK, Dittmann A, Giotopoulos G, Bantscheff M, Chan WI, et al. Inhibition of BET recruitment to chromatin as an effective treatment for MLL-fusion leukaemia. Nature 2011;478:529–33.
49. De Raedt T, Beert E, Pasmant E, Luscan A, Brems H, Ortonne N, et al. PRC2 loss amplifies Ras-driven transcription and confers sensitivity to BRD4-based therapies. Nature 2014;514:247–51.
50. Zuber J, Shi JW, Wang E, Rappaport AR, Herrmann H, Sison EA, et al. RNAi screen identifies Brd4 as a therapeutic target in acute myeloid leukaemia. Nature 2011;478:524–U124.

Published OnlineFirst April 26, 2017; DOI: 10.1158/2159-8290.CD-16-1297

51. Delmore Jake E, Issa Ghayas C, Lemieux Madeleine E, Rahl Peter B, Shi J, Jacobs Hannah M, et al. BET bromodomain inhibition as a therapeutic strategy to target c-Myc. Cell 2011;146:904–17.
52. Rathert P, Roth M, Neumann T, Muerdter F, Roe J-S, Muhar M, et al. Transcriptional plasticity promotes primary and acquired resistance to BET inhibition. Nature 2015;525:543–7.
53. Lockwood WW, Zejnullahu K, Bradner JE, Varmus H. Sensitivity of human lung adenocarcinoma cell lines to targeted inhibition of BET epigenetic signaling proteins. Proc Nat Acad Sci 2012;109:19408–13.
54. Thalhammer V, Lopez-Garcia LA, Herrero-Martin D, Hecker R, Laubscher D, Gierisch ME, et al. PLK1 Phosphorylates PAX3–FOXO1, the inhibition of which triggers regression of alveolar rhabdomyosarcoma. Cancer Res 2015;75:98–110.
55. Ciceri P, Müller S, O'Mahony A, Fedorov O, Filippakopoulos P, Hunt JP, et al. Dual kinase-bromodomain inhibitors for rationally designed polypharmacology. Nat Chem Biol 2014;10:305–12.
56. Linardic CM.PAX3–FOXO1 fusion gene in rhabdomyosarcoma. Cancer Lett 2008;270:10–8.
57. Stewart E, Federico S, Karlstrom A, Shelat A, Sablauer A, Pappo A, et al. The childhood solid tumor network: a new resource for the developmental biology and oncology research communities. Devel Biol 2016;411:287–93.
58. Chang W, Brohl A, Patidar R, Sindiri S, Shern JF, Wei JS, et al. Multidimensional clinomics for precision therapy of children and adolescent young adults with relapsed and refractory cancer: a report from the Center for Cancer Research. Clin Cancer Res 2016;22:3810–20.
59. Barski A, Cuddapah S, Cui K, Roh T-Y, Schones DE, Wang Z, et al. High-resolution profiling of histone methylations in the human genome. Cell 2007;129:823–37.
60. Jin W, Tang Q, Wan M, Cui K, Zhang Y, Ren G, et al. Genome-wide detection of DNase I hypersensitive sites in single cells and FFPE tissue samples. Nature 2015;528:142–6.

Chapter Three: Discussion

Enhancers, the harbor of oncogenic transcription factors and their cofactors

Transcription is commonly deregulated in cancer to support cell proliferation and transformation. Numerous oncogenes encode for transcription factors (TFs) and are recurrently mutated or overexpressed in cancer. For example, chromosomal translocations frequently result in chimeric TFs[230,231] which bind to gene regulatory elements and disrupt gene expression programs.

Gene regulatory elements fall into three basic categories (Fig. 23): promoters, enhancers, and insulators. Promoters lie within 100 to 1000 bp from the TSS (transcription start site) and allow the assembly of basal transcription machinery. Enhancers are distal elements that control promoter activity by recruiting TFs and cofactors. Insulators do not directly influence gene expression, but they delimitate the regions of influence of enhancers[232].

The two pediatric cancers studied above are the result of the activity of two aberrant fusion TFs. Fusion-positive rhabdomyosarcoma (FP-RMS) is driven by PAX3-FOXO1 and Ewing sarcoma (ES) by EWS-FLI1. To further understand how these TFs control gene expression, we need to investigate their genome localization and identify their transcription cofactors.

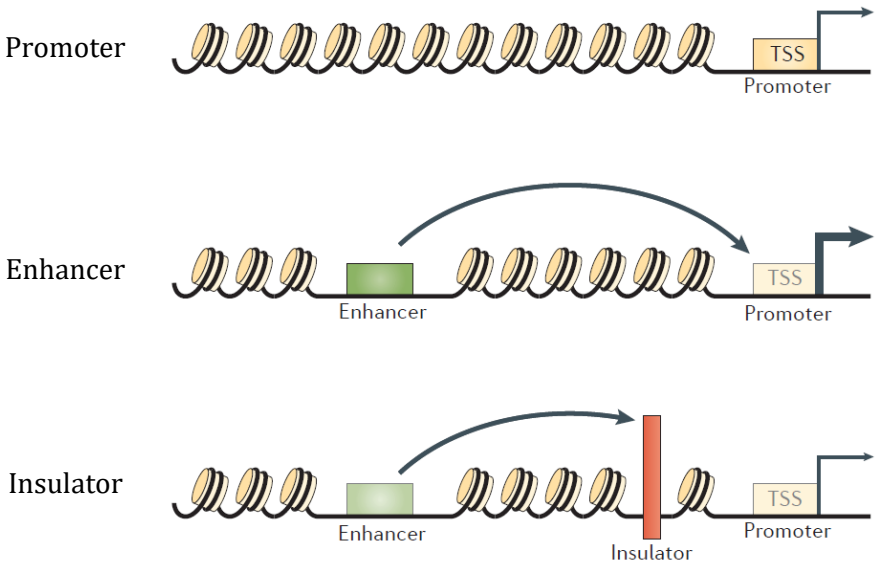

Fig. 23 Gene regulatory elements. Adapted from[232].

1. Characterizing the epigenetic landscape of pediatric sarcomas

Post-translational modifications (PTMs) of histones demarcate the functional elements of the genome (Fig.24). By performing chromatin-immunoprecipitation assays followed by DNA sequencing (ChIP-seq), we can assess the location of the different histone marks and subdivide the chromatin into its different regulatory elements[163]. Then, we can localize TFs and cofactors within the regulatory elements by overlaying their ChIP-seq signal.

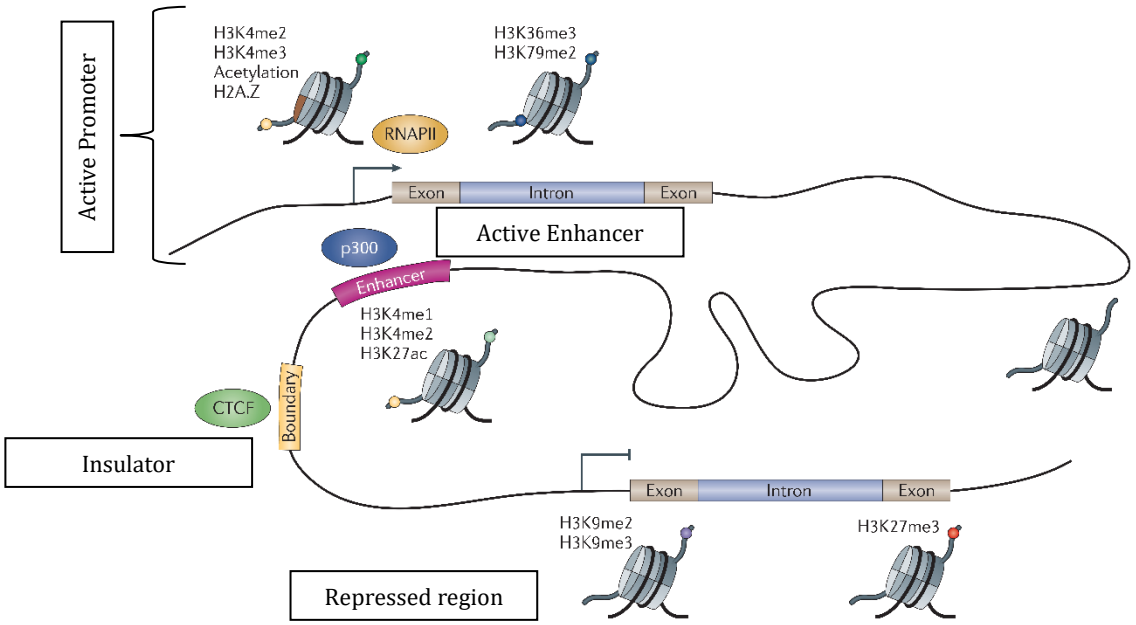

Fig. 24 The differential distribution of histone marks and proteins in the genome and their association with the gene regulatory domains. Adapted from[163].

Our ChIP-seq assays, performed with a monoclonal antibody against PAX3-FOXO1 breakpoint in FP-RMS cell lines, confirmed the previously described localization of PAX3-FOXO1 to distal enhancers[233]. In addition, we observed that most of the super-enhancers (SEs) present in FP-RMS are occupied by the fusion protein.

SEs (Fig.25) are defined as clusters of enhancers that span through large regions of DNA. They are densely decorated by the activation marks H3K27ac and H3K4me1, the mediator complex, the bromodomain-containing protein BRD4, and cell-type-specific TFs. Although the role of SEs as a new element of transcription regulation is still debatable[234], they are usually associated with the control of gene expression related with cell identity, while regular/typical enhancers are associated with the control of genes common to a variety of cell types[235,236]. In agreement with this definition, we found that in FP-RMS SEs are highly occupied by the master transcription factors of myogenesis, MYOD1 and MYOG, as well as by BRD4, p300 and the mediator complex. In addition, most of the SEs found in FP-RMS cells were also co-occupied by the NuRD members HDAC2, RBBP4, and CHD4.

The presence of NuRD in enhancers and super-enhancers has been already described in many cells, such as breast cancer or embryonic stem cells[237–241]. Hence, it has been proposed that rather than being recruited by TFs, NuRD has a general affinity to open chromatin.

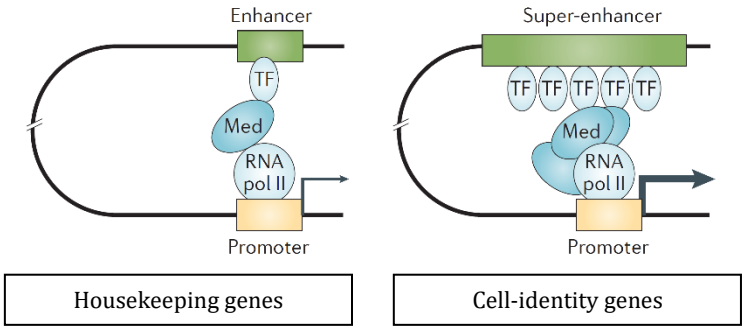

Fig. 25 The structure of typical enhancer versus super-enhancer. Adapted from[232].

In parallel, similar studies performed by others in Ewing sarcoma have identified that the fusion protein EWS-FLI1 binds also to enhancers together with p300 and BRD4. In addition, this fusion protein recruits the chromatin remodeler complex BAF and the component of the MLL complex WDR5 to active enhancers[242–244].

2. Enhancer regulation is essential for tumorigenesis

It is well-known that oncogenic fusion proteins, like PAX3-FOXO1 and EWS-FLI1, are not only unique to their tumor types but also are essential for tumor survival[17]. Transcriptomic studies demonstrate that their relevance in tumorigenesis is caused by the target genes they regulate, which control cell proliferation, invasiveness, and survival[33,100]. Now, a growing interest in epigenetic gene expression regulation has provided clues of how these oncogenic TFs drive their gene expression signatures.

Here, we report that PAX3-FOXO1 fusion establishes new enhancers and super-enhancers together with the master regulators MYOD1, MYOG, and MYCN. In fact, PAX3-FOXO1 depletion reduces the abundance of the enhancer and activation mark H3K27ac and deeply affects enhancer-driven transcription, specially the one of SEs. Consequently, oncogenes, such as ALK and MYCN, are downregulated which leads to tumor cell death. Similarly, in Ewing sarcoma, the fusion protein EWS-FLI1 not only binds to enhancers but also is able to establish de novo enhancers that drive tumorigenesis[242,245]. Unlike PAX3-FOXO1, which mainly activates gene expression, EWS-FLI1 can both activate or repress transcription depending on the underlying DNA sequence. Hence, EWS-FLI1 depletion can both increase or decrease the presence of H3K27ac at enhancers. Activated targets, including CCND1, NKX2-2, EZH2, and SOX2, become repressed and repressed genes, like the known tumor suppressors ERRFI1, CABLES1, IER3, and TGFBI, become activated. These changes in gene expression are then the cause of cell death upon fusion protein silencing.

3. Enhancer targeting as a therapeutic strategy

The dependence of these and other tumors on the expression signatures of their driver oncogenic TFs opens new possibilities for targeted therapy. In fact, targeting transcription regulators, although they are broadly expressed in healthy cells, affects oncogenic expression and cancer cells more than healthy cells[143]. For example, inhibition of cyclin-dependent kinase 7 (CDK7) by THZ1 particularly disrupts MYCN-driven transcription in neuroblastoma cells, and, consequently, inhibits tumor growth[246]. Small cell lung cancer is also sensitive to THZ1 treatment which deeply impairs the expression of the oncogenic TF MYC and other TFs like SOX2, SOX4 and OTX2[247].

SE-associated genes are especially sensitive to perturbations of the transcriptional regulators that populate them. A CRISPR screen in murine embryonic stem cells showed that SEs are dependent on basically all the epigenetic regulators that bind to them to drive gene expression[248]. The loss of the mediator complex in embryonic stem cells preferentially affects genes regulated by SEs and not housekeeping genes, which are controlled by typical enhancers[236]. Pharmacological inhibition of BRD4 also drastically impacts the expression of oncogenes associated with SE and impairs tumor progression in different cancers[249–251]. In multiple myeloma, BRD4 inhibition leads to its displacement from SEs, impaired SE-driven oncogene expression, including MYC, and consequent decrease of cell proliferation[252]. In acute myeloid leukemia (AML), BRD4 depletion removes the mediator complex from enhancers and suppresses the transcription of MYB target genes[250]. Large B-cell lymphoma is also susceptible to BRD4 inhibition which, by diminishing SE-driven transcription, impairs the transcriptional programs of MYC and E2F[251].

Here, we showed that BRD4 depletion caused a preferential decrease of SE-driven gene expression in FP-RMS, as described for other tumors. Moreover, treatment of FP-RMS cell lines with the BRD4 inhibitor JQ1 drastically affected PAX3-FOXO1 target gene signature and impaired tumor growth. Similar findings were observed for CHD4 depletion. Silencing of this chromatin remodeler deeply impaired the SE-driven expression of PAX3-FOXO1 targets. This consequently led to tumor cell death in vitro and tumor regression in vivo. To our knowledge, this is the first report that associates CHD4 with the disruption of SE-associated gene expression in cancer. These findings make both CHD4 and BRD4 possible targets for FP-RMS therapy.

Ewing sarcoma is also dependent on CHD4 and BRD4. Recent data showed that both genetic depletion or pharmacological inhibition of BRD4 impaired enhancer-driven expression of EWS-FLI1 target genes, decreased cell viability and delayed tumor growth[244,253–255]. As observed in FP-RMS, JQ1 treatment removed BRD4 from the chromatin. These results suggest a very similar effect of BRD4 inhibition in FP-RMS and ES.

Since CHD4 was shown to co-regulate the expression of repressed targets of EWS-FLI1[112], we hypothesize that, similarly to FP-RMS, CHD4 is required for EWS-FLI1 gene expression regulation at enhancers. In fact, our results showed that CHD4 depletion causes a drastic decrease in ES cell viability and reduces tumor growth, which makes CHD4 not only a possible therapeutic target in FP-RMS but also in Ewing sarcoma. Currently, we are collecting data to support the function of CHD4 as an EWS-FLI1 coregulator of enhancer driven gene expression. We will perform ChIP-seq assays to identify the genome localization of CHD4. Then, we will correlate EWS-FLI1 location with the one of CHD4 and perform RNA-seq assays to evaluate the effect of this chromatin remodeler in EWS-FLI1 target genes. Since, EWS-FLI1 regulates both gene activation and repression, it will be interesting to know if CHD4 participates in one of these processes or both. Nevertheless, data gathered by us and others support CHD4 and BRD4 as possible therapeutic targets in these two childhood malignancies.

4. CHD4 a regulator of enhancer accessibility

Previously published work has already localized CHD4 to enhancers[240], but the role it plays in those locations has not been clarified yet. Cooperation between CHD4 and transcription factors in gene expression regulation has also been described, although the exact mechanism by which CHD4 contributes to gene expression has not been elucidated. During differentiation of T helper 2 cells, Chd4 forms complexes with the TF Gata3 which are able to either activate or repress gene expression[208]. In glioblastoma, CHD4 interacts with the TF ZFHX4 and regulates gene expression related with stem cell-like functions and tumorigenicity[256]. Here, we reported that CHD4 is able to physically interact with the fusion PAX3-FOXO1 in FP-RMS, while others have described an interaction between CHD4 and EWS-FLI1 in ES[112].

CHD4 is a member of the SNF2 family and has ATP-dependent nucleosome disruption activity similar to the one of the BAF complex[175], meaning that it is able to control nucleosome-spacing[170]. Due to its remodeling function, we wondered if CHD4 functioned by keeping the chromatin open and permissive to the binding of TFs and cofactors to enhancers. DNase hypersensitivity assays demonstrated that CHD4 silencing deeply affects the DNA accessibility at SE which results in a decrease of the ChIP signal of PAX3-FOXO1 and of its cofactor BRD4 at those locations. These results suggest that CHD4 is necessary to keep the chromatin available to the binding of the transcription machinery (Fig.26). Similarly, in mouse embryonic stem cells, CHD4 was also found to regulate the positioning of the TF Sall4[238]. There, the dissolution of the NuRD complex, and consequent loss of CHD4

remodeling activity, also conditioned the access of transcription factors to regulatory sequences[241].

Additionally, we found that CHD4 interacts with BRD4. These results suggest that these proteins might collaborate in the regulation of SE-driven transcription. Interestingly, a

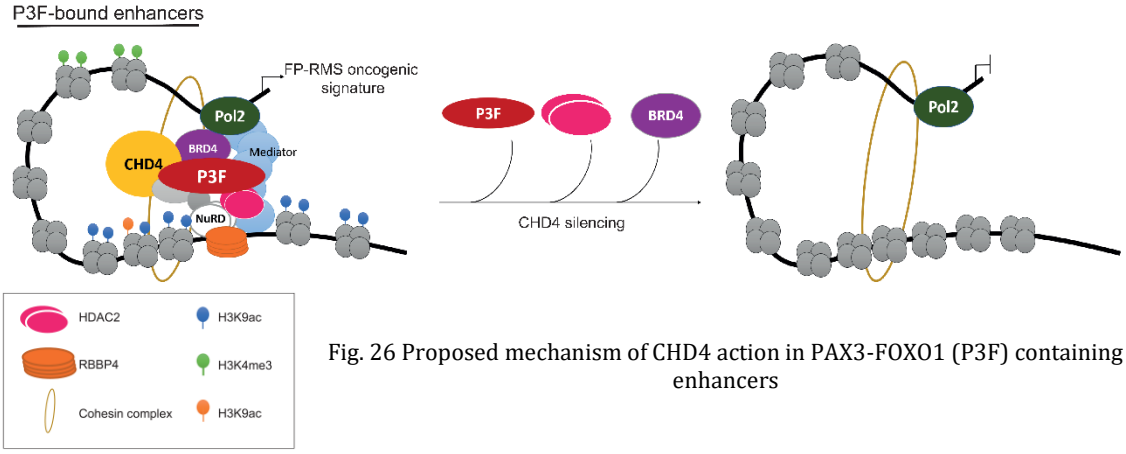

Fig. 26 Proposed mechanism of CHD4 action in PAX3-FOXO1 (P3F) containing enhancers

direct physical interaction between CHD4 and the extra-terminal (ET) domain of BRD3 has already been reported[257]. Due to the high conservation of the ET domain, a similar binding interaction might occur between CHD4 and BRD4.

Since BRD4 exerts partially its function in transcription activation by controlling RNA Pol 2 pause release[258], and since CHD4 interacts with it, we will further investigate if CHD4 plays a role in Pol2 pause release as well. In fact, silencing of Chd4 in mouse embryonic stem cells was shown to decrease Pol2 levels from active promoters[259]. To evaluate the effect of chromatin remodeling in Pol2 positioning in FP-RMS genome, we will track Pol2 location by performing ChIP-seq assays in the presence and absence of CHD4.

In Ewing sarcoma, we observed that CHD4 interacts also with BRD4 in ES cell lines (data not shown). Moreover, BRD4 in ES exists in a large protein complex together with EWS-FLI1, RNA Pol2 and Mediator[244]. These results suggest a similar mechanism of gene regulation between ES and FP-RMS, where CHD4 and BRD4 collaborate to regulate the fusion protein activity.

5. Lessons learnt from CHD4 transcriptomics in FP-RMS

The role of CHD4 in gene expression regulation is so far ambiguous and seems to be dependent on the context, meaning on which cofactors interact with CHD4. In T helper 2 cells, Chd4 activates gene expression together with p300 and represses it together with other members of NuRD[208]. In colon cancer, it collaborates with DNMTs at promoter regions to silence tumor suppressor genes through DNA methylation[260,261]. In mouse embryonic stem cells, Chd4 is required for transcription driven from active promoters while it represses the one from bivalent promoters[259].

In FP-RMS, we found that CHD4 silencing leads to upregulation and downregulation of approximately the same number of genes, suggesting a both activating and repressive role. A study of CHD4 in 293 cells also revealed that it can equally activate and repress gene expression[187]. The changes described above in chromatin architecture and cofactor binding at SE suggest a direct effect of CHD4 on gene activation in FP-RMS. Co-immunoprecipitation assays demonstrated that CHD4 also interacts with the PRC2 members EZH2 and SUZ12 in FP-RMS (data not shown), which could imply that CHD4 might control transcription repression through PRC2. Indeed, during astroglial differentiation Chd4 interacts with Ezh2 to suppress the expression of the key astrogenic marker Gfap[212].

Gene set enrichment analysis showed that the FP-RMS signature was highly affected by CHD4 depletion. Moreover, enhancer-driven gene expression was negatively correlated with CHD4 silencing, which supports the described role of CHD4 in SE regulation.

Genes related with development (e.g. components of the BMP pathway) and myogenesis (e.g. Myl1) were also found among the CHD4 signature. This is in agreement with the growing evidence that shows that CHD4 plays a key role during muscle differentiation and facilitates lineage specification. During muscle differentiation, this chromatin remodeler, through binding to distal regulatory regions, directly regulates the expression of genes required for cardiac muscle development[211]. Embryos without CHD4 can form a normal early blastocyst but are unable to successfully complete the first lineage decision and form trophectoderm, which prevents implantation and leads to embryo lethality[262]. In other cellular contexts, CHD4 has also been associated with cell-fate. During lymphocyte development, the TF Ikaros binds to active regions with CHD4 and to regulates its function. In fact, upon silencing of Ikaros, CHD4 genome location changes which produces alterations in chromatin architecture and decreases the expression of lymphoid cell-specific gene expression which enables the progression of the leukemic state[263].

Surprisingly, CHD4 transcriptomic data revealed that this chromatin remodeler controls the expression of known EWS-FLI1 target genes in FP-RMS. This result allowed us to speculate that the similarities between both ES and FP-RMS go beyond their pediatric presentation. In fact, despite the many efforts to identify the oncogenic signature of these sarcomas, it is still not known which PAX3-FOXO1 or EWS-FLI1 target genes are responsible for cell transformation. The discovery of a common gene signature among these tumors could help to elucidate this question.

The function of the NuRD complex in FP-RMS

The NuRD complex is uniquely characterized by two enzymatic functions: ATP-dependent chromatin remodeling, carried out by CHD3 or CHD4, and histone deacetylase activity,

performed by HDAC1/2. More recently, it has been shown that the lysine-specific histone demethylase 1 (LSD1) is also be associated with the this complex in certain cell types[264]. Other non-enzymatic subunits include methyl-CpG-binding domains (MBD2/3), metastasis-associated proteins (MTA1/2/3), retinoblastoma-binding proteins (RBBP4/7), and GATAD2A/B[201]. Multiple functions have been attributed to NuRD and it is so far thought that combinatorial assembly of the non-enzymatic subunits confers functional specificity.

1. CHD4 dependency on NuRD

Although CHD4 is many times described as a core member of the complex, a recent publication has defined it as a peripheral subunit. In that report, the authors demonstrate the existence of a NuDe complex to which CHD4 can bind to reconstitute the NuRD[265]. In addition, it was already shown that CHD4 can function as an ATPase by itself[195,196].

Our results show that CHD4 co-localizes with PAX3-FOXO1 to enhancers together with the NuRD members HDAC2 and RBBP4. Moreover, BRD4 not only interacts with CHD4 but also with the NuRD members HDAC1/2 and MTA2. Since CHD4 depletion deeply affects FP-RMS cell viability, we hypothesized that if its activity was dependent on NuRD, the dissolution of the complex would also lead to cell death. To test this, we performed a NuRD-centered CRISPR screen and found that knockout of RBBP4, HDAC1, and GATAD2A/B decreased FP-RMS cell proliferation. RBBP4 and HDAC1 are found in other remodeling complexes besides NuRD[178]. Hence, the results from our NuRD screen might reflect the effect of other complexes. Furthermore, unlike CHD4 depletion, silencing of RBBP4 did not affect the expression of PAX3-FOXO1 target genes nor did it cause tumor cell death. The effect on cell viability upon HDAC1 knockout is in agreement with the antiproliferative effect of HDAC inhibitors in FP-RMS described in the literature[266–268]. Nonetheless, these inhibitors seem to act in a ROS-dependent manner and independently of a direct effect on the chromatin[266]. These results support a different function of HDAC1 and RBBP4, in comparison with CHD4, in FP-RMS cell proliferation.

Interestingly, in mouse embryonic stem cells, in the absence of nuclear Mbd3 Chd4 remains only associated with Gatad2b in the nucleus[241]. Moreover, re-establishing nuclear Mbd3 did not affect the chromatin localization of CHD4[241]. These findings support a NuRD-independent function of CHD4 and could explain why we observe an effect upon GATAD2A/B knockout but not with other NuRD members. Moreover, GATAD2A/B, like CHD4, have been described as peripheral members of the NuRD[269].

The NuRD complex is formed by an assembly of binary complexes, one containing the MTA, HDAC, and RBBP subunits and the other with MBD and GATAD2[186]. The MBDs are mutually exclusive in NuRD and commonly referred as the core members of the complex[179,269]. Thus,

the absence of any effect upon knockout of the MTA or MBD proteins suggests that the cell death phenotype observed upon CHD4 silencing in FP-RMS is independent of the assembly of the NuRD complex.

2. A CHD4-free NuRD complex localizes to promoters in FP-RMS

In FP-RMS, the localization of the NuRD components HDAC2 and RBBP4 highly correlates with histone activating marks H3K27ac, H3K4me1, and H3K4me3. This correlation was also observed in mouse embryonic stem cells (ESCs)[241]. We observed that the ChIP peaks of CHD4 and HDAC2+RBBP4 did not overlap completely. This pattern was also observed by others[241] and is consistent with the already described NuRD-independent function of CHD4. Curiously, the CHD4-free HDAC2+RBBP4 regions were highly enriched in the promoter mark H3K4me3 (Fig.27). The absence of CHD4 from promoters is in line with the fact that methylation of H3K4 inhibits the binding of the PHD domains of CHD4 to histones[193,194]. Nevertheless, it is contradictory to other studies where, in other cellular contexts, NuRD-containing CHD4 was shown to localize to promoters as well as enhancers [209,214,241,263,270–272].

Since CHD4 and CHD3 are mutually exclusive subunits of the complex, the promoter regions enriched in HDAC2 and RBBP4 might represent CHD3-containing NuRD locations. In fact, studies have shown that CHD3 and CHD4 have distinct nuclear localization and also carry different remodeling activities[187].

The NuRD occupied promoters regulated the expression of genes related to housekeeping functions such as RNA splicing and ribosome biogenesis. Housekeeping and developmental gene expression are known to be differently regulated[273,274]. Our data suggests that enhancers related with development and muscle differentiation are occupied by CHD4, while promoters involved in cell homeostasis contain CHD4-free NuRD, which can explain the lack of effect in cell proliferation upon NuRD disruption.

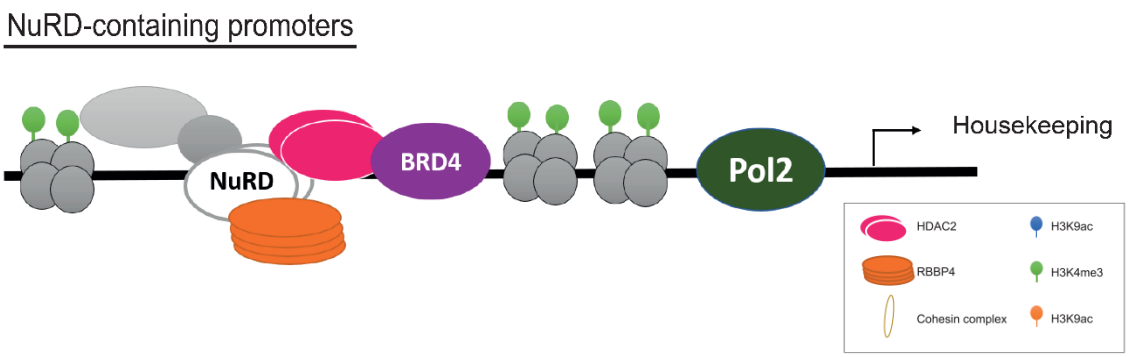

Fig. 27 Scheme of CHD4-free NuRD presence in promoters.

3. NuRD, a repressor and activator

NuRD is classically defined as a repressor, and many published papers corroborate this. At poised genes, NuRD deacetylates H3K27 enabling the recruitment of PRC2 and subsequent gene repression by methylation of H3K27[209]. In human oral squamous cell carcinomas, NuRD opposes the SWI/SNF complex and again deacetylases H3K27 which leads to the binding of PRC2 and KDM1A, followed by H3K27 methylation and H3K4 demethylation[272]. Nevertheless, a gene expression activation capacity has also been attributed to this complex. NuRD collaborates with the TF GATA1 for activation and repression of its target genes[207]. More recently, a role in fine-tuning gene expression has also been described for NuRD. In ESCs, NuRD is found in promoters of actively transcribed genes and is required there to restrict the expression of pluripotency-associated genes, rather than to repress it completely[214]. Hence, NuRD can attenuate the expression of such genes and keeps them responsive to differentiation signals. Another study in mouse ESCs claims that NuRD presence in active enhancers and promoters is essential to fine-tune gene expression response to outside signals and allows proper lineage commitment. Indeed, loss of Mbd3, and consequent dissolution of NuRD, results in only moderate changes in gene expression[241]. Similarly, another study showed that NuRD binds to promoters of rRNA (ribosomal RNA) and keeps them poised but permissive to transcription activation. In that study, knockdown of CHD4 impaired the association of Pol I with rDNA and inhibited rRNA synthesis[275].

In FP-RMS, we found NuRD components in active promoters of housekeeping genes, but it was not clear if the presence of CHD4-free NuRD in those regions acts to restrict transcription, as suggested in previous studies. NuRD components were also present in active enhancers. There, our results show that silencing of RBBP4 did not affect the expression of genes related to those enhancers, suggesting that NuRD does not play an active role there. Nonetheless, we cannot exclude that RBBP4 depletion is perhaps not enough to dissolve the NuRD complex. Thus, studies on other NuRD members are necessary to validate this conclusion.

4. NuRD in Ewing sarcoma

The transcriptional repressive activity of EWS-FLI1 seems to be dependent on NuRD[112]. The fusion protein interacts with the NuRD members HDAC1, CHD4, and MTA2, and HDAC inhibition reduces c-MYC and EWS-FLI1 levels and leads to re-expression of neuronal differentiation markers[276]. In addition, pharmacological inhibition of HDACs reduces ES cell proliferation, impairs colony formation, and induces apotosis[276,277]. In addition, a high-throughput small molecule screen demonstrated that HDAC inhibitors diminish

accessibility at EWS-FLI1 target sites and disrupt the fusion protein-mediated transcription[278]. These results led to the advance of these inhibitors into clinical trials. In fact, phase I trials have already shown that Pracinostat is reasonably well tolerated in children with refractory Ewing sarcoma[279]. Besides direct effect on gene expression, HDAC inhibitors may also work by regulating p53 function. In fact, HDAC1 interacts with p53 and keeps it deacetylated and inactive[280].

Another member of NuRD that affects Ewing sarcoma is LSD1. Pharmacological inhibition of LSD1 affects both activated and repressed EWS-FLI1 target genes[281]. This effect seems to be EWS-FLI1-dependent since LSD1 inhibition in HEK cells had no effect in selected EWS-FLI1 target genes[112]. In addition, LSD1 inhibition decreases ES cell proliferation, induces apoptosis, and reduces tumor growth in a xenograft mouse model[281].

As previously stated, we demonstrated that silencing of CHD4 in Ewing sarcoma leads to cell death and tumor regression, but we do not know if this phenotype is related to NuRD. To investigate this, we are currently performing ChIP-seq assays not only for CHD4 but also for HDAC2 and RBBP4 to identify the location of NuRD in Ewing sarcoma. Moreover, we are performing a similar NuRD-centered CRISPR screen in ES cell lines to analyze the vulnerability of this tumor to the NuRD subunits. Then, we will associate the phenotypical effects observed upon CHD4/NuRD knockout to their genome location and their relationship with the fusion protein EWS-FLI1. At last, we will identify genes coregulated by CHD4 and the fusion by RNA-seq and correlate them to alterations in chromatin accessibility obtained by DNase hypersensitivity assays. These experiments will allow us to understand if the role played by CHD4 in Ewing sarcoma is similar to the one that we described for this remodeler in FP-RMS.

Targeting CHD4

We have shown that CHD4 causes tumor cell death in FP-RMS and Ewing sarcoma. In addition, we showed that CHD4 depletion had no effect on cell proliferation of healthy human myoblasts and fibroblasts. These results were exciting and revealed a tumor specific susceptibility to this chromatin remodeler, which was confirmed by other authors. A RNAi screen identified CHD4 as an essential gene in breast cancer with no effect on cell proliferation of non-transformed mammary epithelial cells (MCF10A)[282]. Another study has shown that CHD4 is required for the maintenance of childhood acute myeloid leukemia but not for the proliferation and survival of normal primary hematopoietic cells[283]. These studies demonstrate the potential of CHD4 as a drug target.

1. Structure-based drug design

CHD4 is composed of a C-terminal ATPase, 2 plant homeodomains (PHD1/2), 2 chromodomains and 2 C-terminal domains of unknown function (DUF)[190]. The chromodomains bind directly to DNA, while the PHD domains interact with histones[191]. High resolution structure information, mainly NMR based, is available for both chromodomains and PHD domains[193,284] but not for the ATPase, most likely due to its high complexity and molecular weight which makes protein purification and NMR studies more difficult. All proteins from the SNF2-like family[166,169,171], which carry a diversity of essential physiological roles, have similar ATPase domains. This would make the design of specific inhibitors targeting the ATPase domain of CHD4 possibly unfeasible.

The ATPase activity of CHD4 is dependent on the simultaneous binding of the PHD domains to histones[195,196] and of the chromodomains to DNA, which suggests that targeting either the PHDs or chromodomains might impair CHD4's remodeling activity. Since protein-DNA interactions are difficult to inhibit with small molecules and inhibitors of histone reader domains have already been successfully developed, as exemplified by the case of the BRD4 bromodomain inhibitor JQ1[226], we decided to target the reader domains of CHD4, more specifically its second PHD domain as it has higher affinity to the histone tail than PHD1[285]. Plant homeodomains (PHDs) are small (circa 65 residues) zinc-finger reader domains that can recognize and bind to different histone marks[286]. The PHD domains of CHD4 recognize the N-tail of histone 3, with increased affinity when Lys9 is methylated or acetylated[192–194]. This domain fold is quite loose, with a short two-stranded antiparallel β-sheet, a short C-terminal α-helix, and three unstructured loops[287]. This minimal fold is maintained by the Zinc ions and conserved hydrophobic residues that form a small hydrophobic core[288]. The histone-binding surface of the PHD finger is shallow and polar. Therefore, while still deemed feasible targets for structure-based drug design efforts, PHD fingers are very difficult targets[289]. Indeed, only a few reports of successful efforts in identifying inhibitors of H3K4-binding PHD fingers are available[290,291].

Our *in silico* efforts gathered a library of 105 candidate PHD2-CHD4 ligands which were validated by NMR binding assays, based on protein and ligand observation. One compound, DSP_C42_003, showed weak binding capability towards the PHD2 domain. Although this ligand should undergo optimization cycles in order to improve its binding affinity, this is the first small molecule binder of a H3K9-binding PHD domain and also the first CHD4 candidate inhibitor identified. We believe that this study represents a first attempt on targeting CHD4 and that it suggests the feasibility of structure-based drug design as a means to develop inhibitors against difficult targets such as PHD domains.

Chromatin remodelers in cancer therapy

Cancer is a very heterogenous disease with a high degree of variability within tumor types, which makes it difficult to find common drug targets. Nevertheless, many studies have been carried out to identify essential genes to which the tumor cell might be vulnerable[292].

BET-bromodomains inhibitors cause drastic decreases in cell proliferation in many cancer types, which has brought attention to epigenetics as a promising field for cancer therapy. Since its development, the BRD4 inhibitor JQ1 has been shown, in vitro and in vivo, to affect many hematological malignancies and a variety of solid tumors such as glioblastoma, medulloblastoma, hepatocellular carcinoma, colon cancer, pancreatic cancer, prostate cancer, lung cancer, and breast cancer[293]. Two phase I clinical trials were performed to test the BET-bromodomain inhibitor OTX015, which targets BRD2, BRD3, and BRD4, in patients with myeloma, lymphoma and acute myeloid leukemia. These trials showed an acceptable toxicity and a good therapeutic response for some patients, although no predictor markers were found[294,295]. These studies show that, despite the broad expression of these bromodomains, we are still able to find therapeutic windows for treatment with epigenetic inhibitors.

Despite the growing evidence that chromatin remodelers play important roles in cancer, they have been so far overlooked as drug targets. Here, we describe in detail the mechanism by which the chromatin remodeler CHD4 influences FP-RMS and provide concrete evidence of its potential as a drug target for both FP-RMS and ES. Other tumors, like synovial sarcoma, also have their tumorigenesis process dependent on chromatin remodeling. Synovial sarcoma is a pediatric malignancy driven by the aberrant fusion protein SS18-SSX[296]. Unlike PAX3-FOXO1 and EWS-FLI1, SS18-SSX does not have a DNA binding sequence. Nevertheless, it is able to reshape gene expression by interacting with the chromatin remodeler BAF complex and reverse Polycomb-mediated gene repression[297,298]. Hence, we decided to take a closer look at chromatin remodelers as potential drug targets.

1. CHD4 as a broad cancer susceptibility

We took advantage of two publicly available databases of tumor susceptibilities, Achilles and DRIVE, to identify cancer sensitivities to chromatin remodelers. Achilles database is a project of the BROAD Institute aimed at identifying essential cancer genes by using genome-scale CRISPR/Cas9 technology to knockout thousands of individual genes in 391 cancer cell lines. Since the Achilles project is still ongoing, the number of guide RNAs used per target varies. The DRIVE is a project of Novartis which shares the same goal as the Achilles, but uses 20 shRNAs per target, instead of CRISPR, in 387 cancer cell lines[299].

The analysis of all SNF2-like ATPases present in chromatin remodelers in those two platforms clearly identified CHD4 as a new tumor susceptibility[189]. We observed that CHD4 depletion, like the one of BRD4, impaired tumor cell viability in virtually all cancer cells analyzed in both projects (in the DRIVE database 323 out of 387 and in the Achilles database 386 out of 391 cancer cell lines). This astonishing result is in line with published data where CHD4 silencing was shown to affect breast cancer[282], glioblastoma[300], acute myeloid leukemia[283], liver cancer[301], lung cancer[302], and colorectal cancer[260,261]. Furthermore, CHD4 is rarely found mutated in cancer[216], which is an indicator of tumor dependency. Nevertheless, a missense single nucleotide variation (SNV) found in the N-terminal of CHD4 was associated with lung cancer, malignant lymphoma, rectum cancer, and to interact with smoking in order to confer it a higher risk. Of note, this SNV did not affect neither the ATPase nor the PHD domains of CHD4, and its implications in CHD4 function are yet to be described[303]. In addition, in colorectal cancer and hepatocellular carcinoma, CHD4 levels are correlated with a poorer prognosis[260,301]. In glioma, amplification of the CHD4 gene, which results in CHD4 overexpression, is also associated with an increase in the recurrence rate[304]. A correlation with prognosis was not detected by our studies in FP-RMS.

The mechanism of dependence to CHD4 may vary in the different tumors. Colorectal cancer depends on CHD4 since it mediates the silencing of tumor suppressor genes by recruitment of DNMTs and EZH2[260]. In glioblastoma, CHD4 permits the transcription driven by the tumor-associated TF ZFHX4[256]. In acute myeloid leukemia, CHD4 depletion induces transcription alterations that impair colony formation[305]. Here, we have revealed a transcriptional addiction of FP-RMS to CHD4. Besides the implications of CHD4 in transcription, other functions might be also related to its vast tumor susceptibility. For example, CHD4 has been highly associated with DNA-damage repair. In fact, CHD4 depletion leads to the accumulation of spontaneous DNA damage, increased ionizing radiation sensitivity, and impaired homologous recombination repair[198,199].

CHD4 inhibition could also be used in combination therapy. In hepatocellular carcinoma, CHD4 knockdown increases chemosensitivity to epirubicin[301] and, in acute myeloid leukemia, it leads to a global relaxation of chromatin which renders the cell more susceptible to double-strand breaks and to the genotoxic action of agents such as daunorubicin and cytarabine[305].

In conclusion (Fig.28), we have identified a new general tumor dependency to the chromatin remodeler CHD4. Importantly, we and others have shown that healthy cells are not susceptible to CHD4 inhibition, which makes it a promising target for therapy. In addition, our studies in FP-RMS revealed a new function of CHD4 in the regulation of SE-driven oncogene expression which might explain the vulnerability of transcriptionally addicted cancers to CHD4. At last, we initiated the development of the first CHD4 inhibitor by studying the feasibility of targeting one of its PHD reader domains. This work puts CHD4 side-by-side with BRD4 as a prominent tumor target and hopefully will motivate the drug discovery field to look at CHD4 as promising new cancer target.

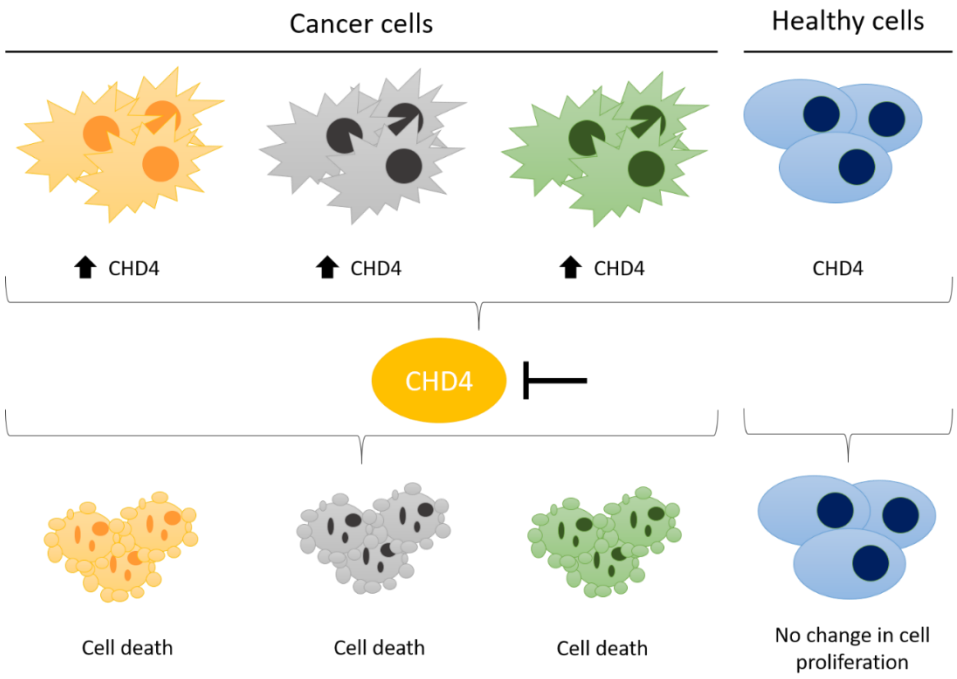

Fig. 28 Tumor susceptibility to CHD4. Many tumor types overexpress CHD4 in comparison to healthy cells. CHD4 depletion seems to not affect the proliferation of the healthy cells but to cause tumor cell death, possibly by apoptosis.

Chapter Four: References and Acknowledgments

References

1. Early History of Cancer | American Cancer Society. Available at: https://www.cancer.org/cancer/cancer-basics/history-of-cancer/what-is-cancer.html. (Accessed: 15th January 2018)

2. Sudhakar, A. History of Cancer, Ancient and Modern Treatment Methods. *J. Cancer Sci. Ther.* **1,** 1–4 (2009).

3. cancer | Definition of cancer in English by Oxford Dictionaries. Available at: https://en.oxforddictionaries.com/definition/cancer. (Accessed: 15th January 2018)

4. Data and statistics. (2018). Available at: http://www.euro.who.int/en/health-topics/noncommunicable-diseases/cancer/data-and-statistics. (Accessed: 15th January 2018)

5. Leading causes of death in Europe: fact sheet. Available at: http://www.euro.who.int/__data/assets/pdf_file/0004/185215/Leading-causes-of-death-in-Europe-Fact-Sheet.pdf?ua=1. (Accessed: 15th January 2018)

6. Siegel, R. Cáncer Statistics. *Ca Cáncer J.* **67,** 7–30 (2017).

7. Ferlay, J. *et al.* Cancer incidence and mortality patterns in Europe: Estimates for 40 countries in 2012. *Eur. J. Cancer* **49,** 1374–1403 (2013).

8. Sonnenschein, C. & Soto, A. M. Theories of carcinogenesis: an emerging perspective. *Semin. Cancer Biol.* **18,** 372–7 (2008).

9. Davidoff, A. M. Pediatric oncology. *Semin. Pediatr. Surg.* **19,** 225–33 (2010).

10. Petrucelli, N., Daly, M. B. & Feldman, G. L. Hereditary breast and ovarian cancer due to mutations in BRCA1 and BRCA2. *Genet. Med.* **12,** 245–259 (2010).

11. Kim, S. New and emerging factors in tumorigenesis: an overview. *Cancer Manag. Res.* **7,** 225 (2015).

12. Hanahan, D. & Weinberg, R. A. Hallmarks of cancer: the next generation. *Cell* **144,** 646–74 (2011).

13. Vogelstein, B. *et al.* Cancer genome landscapes. *Science* **339,** 1546–58 (2013).

14. Robert Weinberg. *The Biology of Cancer.* (2013). doi:10.1016/0160-9327(82)90105-3

15. Arruebo, M. *et al.* Assessment of the evolution of cancer treatment therapies. *Cancers (Basel).* **3,** 3279–330 (2011).

16. Cancer survival statistics for all cancers combined | Cancer Research UK. Available at: http://www.cancerresearchuk.org/health-professional/cancer-statistics/survival/all-cancers-combined#heading-One. (Accessed: 19th June 2018)

17. Dupain, C., Harttrampf, A. C., Urbinati, G., Geoerger, B. & Massaad-Massade, L. Relevance of Fusion Genes in Pediatric Cancers: Toward Precision Medicine. *Mol. Ther. Nucleic Acids* **6,** 315–326 (2017).

18. Ward, E., DeSantis, C., Robbins, A., Kohler, B. & Jemal, A. Childhood and adolescent cancer statistics, 2014. *CA. Cancer J. Clin.* **64,** 83–103 (2014).

19. Downing, J. R. *et al.* The Pediatric Cancer Genome Project. *Nat. Genet.* **44,** 619–622 (2012).

20. Howlader N, Noone AM, Krapcho M, Miller D, Bishop K, Kosary CL, Yu M, Ruhl J, Tatalovich Z, Mariotto A, Lewis DR, Chen HS, Feuer EJ, C. K. Cancer Statistics Review, 1975-2014 - SEER Statistics. *National Cancer Institute. Bethesda, MD* (2017). Available at: https://seer.cancer.gov/csr/1975_2014/. (Accessed: 16th January 2018)

21. Pui, C.-H., Gajjar, A. J., Kane, J. R., Qaddoumi, I. A. & Pappo, A. S. Challenging issues in pediatric oncology. *Nat. Rev. Clin. Oncol.* **8,** 540–549 (2011).

22. Chen, X., Pappo, A. & Dyer, M. A. Pediatric solid tumor genomics and developmental pliancy. *Oncogene* **34,** 5207–5215 (2015).

23. Anderson, J. L., Denny, C. T., Tap, W. D. & Federman, N. Pediatric sarcomas: translating molecular pathogenesis of disease to novel therapeutic possibilities. *Pediatr. Res.* **72,** 112–121 (2012).

24. De Giovanni, C., Landuzzi, L., Nicoletti, G., Lollini, P.-L. & Nanni, P. Molecular and cellular biology of rhabdomyosarcoma. *Futur. Oncol.* **5,** 1449–1475 (2009).

25. Marshall, A. D. & Grosveld, G. C. Alveolar rhabdomyosarcoma - The molecular drivers of PAX3/7-FOXO1-induced tumorigenesis. *Skelet. Muscle* **2,** 25 (2012).

26. Huh, W. & Egas Bejar, D. Rhabdomyosarcoma in adolescent and young adult patients: current perspectives. *Adolesc. Health. Med. Ther.* **5,** 115 (2014).

27. Panda, S. P. *et al.* Diagnosis and Management of Rhabdomyosarcoma in Children and Adolescents: ICMR Consensus Document. *Indian J. Pediatr.* **84,** 393–402 (2017).

28. Rudzinski, E. R. *et al.* The World Health Organization Classification of Skeletal Muscle Tumors in Pediatric Rhabdomyosarcoma: A Report From the Children's Oncology Group. *Arch. Pathol. Lab. Med.* **139,** 1281–7 (2015).

29. Kashi, V. P., Hatley, M. E. & Galindo, R. L. Probing for a deeper understanding of rhabdomyosarcoma: insights from complementary model systems. *Nat. Rev. Cancer* **15,** 426–439 (2015).

30. Shern, J. F. *et al.* Comprehensive Genomic Analysis of Rhabdomyosarcoma Reveals a Landscape of Alterations Affecting a Common Genetic Axis in Fusion-Positive and Fusion-Negative Tumors. *Cancer Discov.* **4,** 216–231 (2014).

31. Williamson, D. *et al.* Fusion Gene–Negative Alveolar Rhabdomyosarcoma Is Clinically and Molecularly Indistinguishable From Embryonal Rhabdomyosarcoma. *J. Clin. Oncol.* **28,** 2151–2158 (2010).

32. Wachtel, M. *et al.* Gene Expression Signatures Identify Rhabdomyosarcoma Subtypes and Detect a Novel t(2;2)(q35;p23) Translocation Fusing PAX3 to NCOA1. *Cancer Res.* **64,** 5539–5545 (2004).

33. Davicioni, E. *et al.* Identification of a PAX-FKHR Gene Expression Signature that

Defines Molecular Classes and Determines the Prognosis of Alveolar Rhabdomyosarcomas. *Cancer Res.* **66,** 6936–6946 (2006).

34. Laé, M. *et al.* Global gene expression profiling of PAX-FKHR fusion-positive alveolar and PAX-FKHR fusion-negative embryonal rhabdomyosarcomas. *J. Pathol.* **212,** 143–151 (2007).

35. Hettmer, S. *et al.* Rhabdomyosarcoma: Current Challenges and Their Implications for Developing Therapies. *Cold Spring Harb. Perspect. Med.* **4,** a025650–a025650 (2014).

36. Anderson, J., Ramsay, A., Gould, S. & Pritchard-Jones, K. PAX3-FKHR Induces Morphological Change and Enhances Cellular Proliferation and Invasion in Rhabdomyosarcoma. *Am. J. Pathol.* **159,** 1089–1096 (2001).

37. Shern, J. F., Yohe, M. E. & Khan, J. Pediatric Rhabdomyosarcoma. *Crit. Rev. Oncog.* **20,** 227–43 (2015).

38. Mercado, G. E. & Barr, F. G. Fusions involving PAX and FOX genes in the molecular pathogenesis of alveolar rhabdomyosarcoma: recent advances. *Curr. Mol. Med.* **7,** 47–61 (2007).

39. Gallego, S. *et al.* Molecular Genetics of Rhabdomyosarcoma. in *eLS* (John Wiley & Sons, Ltd, 2012). doi:10.1002/9780470015902.a0023883

40. Missiaglia, E. *et al. PAX3/FOXO1* Fusion Gene Status Is the Key Prognostic Molecular Marker in Rhabdomyosarcoma and Significantly Improves Current Risk Stratification. *J. Clin. Oncol.* **30,** 1670–1677 (2012).

41. Mahajan, P., Leavey, P. J. & Galindo, R. L. PAX genes in childhood oncogenesis: developmental biology gone awry? *Oncogene* **34,** 2681–2689 (2015).

42. Lam, P. Y., Sublett, J. E., Hollenbach, A. D. & Roussel, M. F. The oncogenic potential of the Pax3-FKHR fusion protein requires the Pax3 homeodomain recognition helix but not the Pax3 paired-box DNA binding domain. *Mol. Cell. Biol.* **19,** 594–601 (1999).

43. Scheidler, S., Fredericks, W. J., Rauscher, F. J., Barr, F. G. & Vogt, P. K. The hybrid PAX3-FKHR fusion protein of alveolar rhabdomyosarcoma transforms fibroblasts in culture. *Proc. Natl. Acad. Sci. U. S. A.* **93,** 9805–9 (1996).

44. Fredericks, W. J., Ayyanathan, K., Herlyn, M., Friedman, J. R. & Rauscher, F. J. An engineered PAX3-KRAB transcriptional repressor inhibits the malignant phenotype of alveolar rhabdomyosarcoma cells harboring the endogenous PAX3-FKHR oncogene. *Mol. Cell. Biol.* **20,** 5019–31 (2000).

45. Charytonowicz, E., Cordon-Cardo, C., Matushansky, I. & Ziman, M. Alveolar rhabdomyosarcoma: Is the cell of origin a mesenchymal stem cell? *Cancer Lett.* **279,** 126–136 (2009).

46. Hollenbach, A. D., Sublett, J. E., McPherson, C. J. & Grosveld, G. The Pax3-FKHR oncoprotein is unresponsive to the Pax3-associated repressor hDaxx. *EMBO J.* **18,** 3702–3711 (1999).

47. Begum, S. *et al.* Cell-type-specific regulation of distinct sets of gene targets by Pax3 and Pax3/FKHR. *Oncogene* **24,** 1860–1872 (2005).

48. Anderson, M. J., Shelton, G. D., Cavenee, W. K. & Arden, K. C. Embryonic expression of the tumor-associated PAX3-FKHR fusion protein interferes with the developmental functions of Pax3. *Proc. Natl. Acad. Sci. U. S. A.* **98,** 1589–94 (2001).

49. Hingorani, P. & Kolb, E. A. Past, present and future of therapies in pediatric sarcomas. *Futur. Oncol.* **6,** 605–618 (2010).

50. Margue, C. M., Bernasconi, M., Barr, F. G. & Schäfer, B. W. Transcriptional modulation of the anti-apoptotic protein BCL-XL by the paired box transcription factors PAX3 and PAX3/FKHR. *Oncogene* **19,** 2921–2929 (2000).

51. Mercado, G. E. *et al.* Identification of PAX3-FKHR-regulated genes differentially expressed between alveolar and embryonal rhabdomyosarcoma: Focus onMYCN as a biologically relevant target. *Genes, Chromosom. Cancer* **47,** 510–520 (2008).

52. Graf Finckenstein, F., Shahbazian, V., Davicioni, E., Ren, Y.-X. & Anderson, M. J. PAX-FKHR function as pangenes by simultaneously inducing and inhibiting myogenesis. *Oncogene* **27,** 2004–2014 (2008).

53. Sun, X. *et al.* Rhabdomyosarcoma: Advances in Molecular and Cellular Biology. *Sarcoma* **2015,** 1–14 (2015).

54. Abraham, J. *et al.* Lineage of origin in rhabdomyosarcoma informs pharmacological response. *Genes Dev.* **28,** 1578–1591 (2014).

55. Cancer Epidemiology in Older Adolescents and Young Adults 15 to 29 Years of Age. Available at: https://seer.cancer.gov/archive/publications/aya/aya_mono_complete.pdf. (Accessed: 23rd May 2018)

56. Ozaki, T. Diagnosis and treatment of Ewing sarcoma of the bone: a review article. *J. Orthop. Sci.* **20,** 250 (2015).

57. Ewing, J. Classics in oncology. Diffuse endothelioma of bone. James Ewing. Proceedings of the New York Pathological Society, 1921. *CA. Cancer J. Clin.* **22,** 95–8

58. Delattre, O. *et al.* Gene fusion with an ETS DNA-binding domain caused by chromosome translocation in human tumours. *Nature* **359,** 162–165 (1992).

59. Stiller, C. A. *et al.* Descriptive epidemiology of sarcomas in Europe: Report from the RARECARE project. *Eur. J. Cancer* **49,** 684–695 (2013).

60. Cheung, C. C., Kandel, R. A., Bell, R. S., Mathews, R. E. & Ghazarian, D. M. Extraskeletal Ewing sarcoma in a 77-year-old woman. *Arch. Pathol. Lab. Med.* **125,** 1358–60 (2001).

61. Jawad, M. U. *et al.* Ewing sarcoma demonstrates racial disparities in incidence-related and sex-related differences in outcome. *Cancer* **115,** 3526–3536 (2009).

62. Monument, M. J. *et al.* Clinical and Biochemical Function of Polymorphic NR0B1 GGAA-Microsatellites in Ewing Sarcoma: A Report from the Children's Oncology

Group. *PLoS One* **9,** e104378 (2014).

63. Beck, R. *et al.* EWS/FLI-responsive GGAA microsatellites exhibit polymorphic differences between European and African populations. *Cancer Genet.* **205,** 304–312 (2012).

64. Bernstein, M. *et al.* Ewing's Sarcoma Family of Tumors: Current Management. *Oncologist* **11,** 503–519 (2006).

65. Choi, E.-Y. K., Gardner, J. M., Lucas, D. R., McHugh, J. B. & Patel, R. M. Ewing sarcoma. *Semin. Diagn. Pathol.* **31,** 39–47 (2014).

66. Applebaum, M. A. *et al.* Clinical features and outcomes in patients with extraskeletal ewing sarcoma. *Cancer* **117,** 3027–3032 (2011).

67. Widhe, B. & Widhe, T. Initial symptoms and clinical features in osteosarcoma and Ewing sarcoma. *J. Bone Joint Surg. Am.* **82,** 667–74 (2000).

68. Paulussen, M., Bielack, S., Jurgens, H. & Casali, P. G. Ewing's sarcoma of the bone: ESMO Clinical Recommendations for diagnosis, treatment and follow-up. *Ann. Oncol.* **20,** iv140-iv142 (2009).

69. Delattre, O. *et al.* The Ewing Family of Tumors -- A Subgroup of Small-Round-Cell Tumors Defined by Specific Chimeric Transcripts. *N. Engl. J. Med.* **331,** 294–299 (1994).

70. Doyle, L. A. Sarcoma classification: An update based on the 2013 World Health Organization Classification of Tumors of Soft Tissue and Bone. *Cancer* **120,** 1763–1774 (2014).

71. Brohl, A. S. *et al.* The Genomic Landscape of the Ewing Sarcoma Family of Tumors Reveals Recurrent STAG2 Mutation. *PLoS Genet.* **10,** e1004475 (2014).

72. Crompton, B. D. *et al.* The Genomic Landscape of Pediatric Ewing Sarcoma. *Cancer Discov.* **4,** 1326–1341 (2014).

73. Sand, L. G. L., Szuhai, K. & Hogendoorn, P. C. W. Sequencing Overview of Ewing Sarcoma: A Journey across Genomic, Epigenomic and Transcriptomic Landscapes. *Int. J. Mol. Sci.* **16,** 16176–215 (2015).

74. Owen, L. A. & Lessnick, S. L. Identification of Target Genes in Their Native Cellular Context. *Cell Cycle* **5,** 2049–2053 (2006).

75. Tirode, F. *et al.* Genomic landscape of Ewing sarcoma defines an aggressive subtype with co-association of STAG2 and TP53 mutations. *Cancer Discov.* **4,** 1342–53 (2014).

76. Sankar, S. & Lessnick, S. L. Promiscuous partnerships in Ewing's sarcoma. *Cancer Genet.* **204,** 351–365 (2011).

77. Sorensen, P. H. B. *et al.* A second Ewing's sarcoma translocation, t(21;22), fuses the EWS gene to another ETS–family transcription factor, ERG. *Nat. Genet.* **6,** 146–151 (1994).

78. Jeon, I. S. *et al.* A variant Ewing's sarcoma translocation (7;22) fuses the EWS gene

to the ETS gene ETV1. *Oncogene* **10,** 1229–34 (1995).

79. Kaneko, Y. *et al.* Fusion of anETS-family gene,EIAF, toEWS by t(17;22)(q12;q12) chromosome translocation in an undifferentiated sarcoma of infancy. *Genes, Chromosom. Cancer* **15,** 115–121 (1996).

80. Peter, M. *et al.* A new member of the ETS family fused to EWS in Ewing tumors. *Oncogene* **14,** 1159–1164 (1997).

81. Yustein, J. T. & Virani Ewing, F. D. *EWS-FLI1 Regulates Genotoxic Stress Response in Ewing Sarcoma*. *Journal of Cancer Biology & Research Cite this article: Gargi G, Yustein JT* **3,** (2015).

82. Braunreiter, C. L., Hancock, J. D., Coffin, C., Boucher, K. M. & Lessnick, S. L. Expression of EWS-ETS Fusions in NIH3T3 Cells Reveals Significant Differences to Ewing's Sarcoma. *Cell Cycle* **5,** 2753–2759 (2006).

83. Thompson, A. D., Teitell, M. A., Arvand, A. & Denny, C. T. Divergent Ewing's sarcoma EWS/ETS fusions confer a common tumorigenic phenotype on NIH3T3 cells. *Oncogene* **18,** 5506–5513 (1999).

84. May, W. A. *et al.* Ewing sarcoma 11;22 translocation produces a chimeric transcription factor that requires the DNA-binding domain encoded by FLI1 for transformation. *Proc. Natl. Acad. Sci. U. S. A.* **90,** 5752–6 (1993).

85. Arvand, A., Welford, S. M., Teitell, M. A. & Denny, C. T. The COOH-terminal domain of FLI-1 is necessary for full tumorigenesis and transcriptional modulation by EWS/FLI-1. *Cancer Res.* **61,** 5311–7 (2001).

86. Shing, D. C. *et al.* FUS/ERG gene fusions in Ewing's tumors. *Cancer Res.* **63,** 4568–76 (2003).

87. Ng, T. L. *et al.* Ewing sarcoma with novel translocation t(2;16) producing an in-frame fusion of FUS and FEV. *J. Mol. Diagn.* **9,** 459–63 (2007).

88. May, W. A. *et al.* The Ewing's sarcoma EWS/FLI-1 fusion gene encodes a more potent transcriptional activator and is a more powerful transforming gene than FLI-1. *Mol. Cell. Biol.* **13,** 7393–8 (1993).

89. Prieur, A., Tirode, F., Cohen, P. & Delattre, O. EWS/FLI-1 Silencing and Gene Profiling of Ewing Cells Reveal Downstream Oncogenic Pathways and a Crucial Role for Repression of Insulin-Like Growth Factor Binding Protein 3. *Mol. Cell. Biol.* **24,** 7275–7283 (2004).

90. Kovar, H. *et al.* EWS/FLI-1 antagonists induce growth inhibition of Ewing tumor cells in vitro. *Cell Growth Differ.* **7,** 429–37 (1996).

91. Tanaka, K., Iwakuma, T., Harimaya, K., Sato, H. & Iwamoto, Y. EWS-Fli1 Antisense Oligodeoxynucleotide Inhibits Proliferation of Human Ewing's Sarcoma and Primitive Neuroectodermal Tumor Cells. *J. Clin. Invest* **99,** 239–247 (1997).

92. Fisher, C. The diversity of soft tissue tumours with *EWSR1* gene rearrangements: a review. *Histopathology* **64,** 134–150 (2014).

93. Wei, G.-H. *et al.* Genome-wide analysis of ETS-family DNA-binding in vitro and in vivo. *EMBO J.* **29,** 2147–2160 (2010).

94. Lessnick, S. L., Dacwag, C. S. & Golub, T. R. The Ewing's sarcoma oncoprotein EWS/FLI induces a p53-dependent growth arrest in primary human fibroblasts. *Cancer Cell* **1,** 393–401 (2002).

95. Hu-Lieskovan, S. *et al.* EWS-FLI1 Fusion Protein Up-regulates Critical Genes in Neural Crest Development and Is Responsible for the Observed Phenotype of Ewing's Family of Tumors. *Cancer Res.* **65,** 4633–4644 (2005).

96. Riggi, N. *et al.* EWS-FLI-1 expression triggers a Ewing's sarcoma initiation program in primary human mesenchymal stem cells. *Cancer Res.* **68,** 2176–85 (2008).

97. Smith, R. *et al.* Expression profiling of EWS/FLI identifies NKX2.2 as a critical target gene in Ewing's sarcoma. *Cancer Cell* **9,** 405–16 (2006).

98. Kinsey, M., Smith, R. & Lessnick, S. L. NR0B1 Is Required for the Oncogenic Phenotype Mediated by EWS/FLI in Ewing's Sarcoma. *Mol. Cancer Res.* **4,** 851–859 (2006).

99. Tirado, O. M. *et al.* Caveolin-1 (CAV1) Is a Target of EWS/FLI-1 and a Key Determinant of the Oncogenic Phenotype and Tumorigenicity of Ewing's Sarcoma Cells. *Cancer Res.* **66,** 9937–9947 (2006).

100. Prieur, A., Tirode, F., Cohen, P. & Delattre, O. EWS/FLI-1 Silencing and Gene Profiling of Ewing Cells Reveal Downstream Oncogenic Pathways and a Crucial Role for Repression of Insulin-Like Growth Factor Binding Protein 3. *Mol. Cell. Biol.* **24,** 7275–7283 (2004).

101. Hancock, J. D. & Lessnick, S. L. A transcriptional profiling meta-analysis reveals a core EWS-FLI gene expression signature. *Cell Cycle* **7,** 250–256 (2008).

102. Abaan, O. D. *et al.* PTPL1 is a direct transcriptional target of EWS-FLI1 and modulates Ewing's Sarcoma tumorigenesis. *Oncogene* **24,** 2715–2722 (2005).

103. Park, H.-R. *et al.* Upregulation of the oncogenic helix-loop-helix protein Id2 in Ewing sarcoma. *Tumori* **92,** 236–40

104. Zwerner, J. P. *et al.* The EWS/FLI1 oncogenic transcription factor deregulates GLI1. *Oncogene* **27,** 3282–3291 (2008).

105. Matsumoto, Y. *et al.* Downregulation and forced expression of EWS-Fli1 fusion gene results in changes in the expression of G1regulatory genes. *Br. J. Cancer* **84,** 768–775 (2001).

106. Luo, W. *et al.* GSTM4 is a microsatellite-containing EWS/FLI target involved in Ewing's sarcoma oncogenesis and therapeutic resistance. *Oncogene* **28,** 4126–4132 (2009).

107. Takahashi, A. *et al.* EWS/ETS fusions activate telomerase in Ewing's tumors. *Cancer Res.* **63,** 8338–44 (2003).

108. Yang, L., Hu, H.-M., Zielinska-Kwiatkowska, A. & Chansky, H. A. FOXO1 is a

direct target of EWS-Fli1 oncogenic fusion protein in Ewing's sarcoma cells. *Biochem. Biophys. Res. Commun.* **402,** 129–34 (2010).

109. Nakatani, F. *et al.* Identification of p21 $^{WAF1/CIP1}$ as a Direct Target of EWS-Fli1 Oncogenic Fusion Protein. *J. Biol. Chem.* **278,** 15105–15115 (2003).

110. Hahm, K.-B. *et al.* Repression of the gene encoding the TGF-β type II receptor is a major target of the EWS-FLI1 oncoprotein. *Nat. Genet.* **23,** 222–227 (1999).

111. Boro, A. *et al.* Small-molecule screen identifies modulators of EWS/FLI1 target gene expression and cell survival in Ewing's sarcoma. *Int. J. Cancer* **131,** 2153–2164 (2012).

112. Sankar, S. *et al.* Mechanism and relevance of EWS/FLI-mediated transcriptional repression in Ewing sarcoma. *Oncogene* **32,** 5089–100 (2013).

113. Brenner, J. C. *et al.* PARP-1 Inhibition as a Targeted Strategy to Treat Ewing's Sarcoma. *Cancer Res.* **72,** 1608–1613 (2012).

114. Stewart, E. *et al.* Targeting the DNA Repair Pathway in Ewing Sarcoma. *Cell Rep.* **9,** 829 (2014).

115. Choy, E. *et al.* Phase II study of olaparib in patients with refractory Ewing sarcoma following failure of standard chemotherapy. *BMC Cancer* **14,** 813 (2014).

116. Toretsky, J. A. *et al.* Oncoprotein EWS-FLI1 Activity Is Enhanced by RNA Helicase A. *Cancer Res.* **66,** 5574–5581 (2006).

117. Erkizan, H. V *et al.* A small molecule blocking oncogenic protein EWS-FLI1 interaction with RNA helicase A inhibits growth of Ewing's sarcoma. *Nat. Med.* **15,** 750–756 (2009).

118. Herrero-Martín, D. *et al.* Stable interference of EWS–FLI1 in an Ewing sarcoma cell line impairs IGF-1/IGF-1R signalling and reveals TOPK as a new target. *Br. J. Cancer* **101,** 80–90 (2009).

119. Ackermann, M., Morse, B. A., Delventhal, V., Carvajal, I. M. & Konerding, M. A. Anti-VEGFR2 and anti-IGF-1R-Adnectins inhibit Ewing's sarcoma A673-xenograft growth and normalize tumor vascular architecture. *Angiogenesis* **15,** 685–695 (2012).

120. Kurmasheva, R. T. *et al.* The Insulin-like Growth Factor-1 Receptor-Targeting Antibody, CP-751,871, Suppresses Tumor-Derived VEGF and Synergizes with Rapamycin in Models of Childhood Sarcoma. *Cancer Res.* **69,** 7662–7671 (2009).

121. Jiang, Y., Ludwig, J. & Janku, F. Targeted therapies for advanced Ewing sarcoma family of tumors. *Cancer Treat. Rev.* **41,** 391–400 (2015).

122. Chen, C., Wonsey, D. R., Lemieux, M. E. & Kung, A. L. Differential Disruption of EWS-FLI1 Binding by DNA-Binding Agents. *PLoS One* **8,** e69714 (2013).

123. Grohar, P. J. *et al.* Ecteinascidin 743 interferes with the activity of EWS-FLI1 in Ewing sarcoma cells. *Neoplasia* **13,** 145–53 (2011).

124. Le Deley, M.-C. *et al.* Impact of EWS-ETS fusion type on disease progression in

Ewing's sarcoma/peripheral primitive neuroectodermal tumor: prospective results from the cooperative Euro-E.W.I.N.G. 99 trial. *J. Clin. Oncol.* **28,** 1982–8 (2010).

125. Tu, J. *et al.* The Histogenesis of Ewing Sarcoma. *Cancer reports Rev.* **1,** (2017).

126. Franchi, A. *et al.* Immunohistochemical and ultrastructural investigation of neural differentiation in Ewing sarcoma/PNET of bone and soft tissues. *Ultrastruct. Pathol.* **25,** 219–25

127. Suh, C.-H., Ordóñez, N. G., Hicks, J. & Mackay, B. Ultrastructure of the Ewing's Sarcoma Family of Tumors. *Ultrastruct. Pathol.* **26,** 67–76 (2002).

128. Cavazzana, A. O., Miser, J. S., Jefferson, J. & Triche, T. J. Experimental evidence for a neural origin of Ewing's sarcoma of bone. *Am. J. Pathol.* **127,** 507–18 (1987).

129. von Levetzow, C. *et al.* Modeling Initiation of Ewing Sarcoma in Human Neural Crest Cells. *PLoS One* **6,** e19305 (2011).

130. Lessnick, S. L., Dacwag, C. S. & Golub, T. R. The Ewing's sarcoma oncoprotein EWS/FLI induces a p53-dependent growth arrest in primary human fibroblasts. *Cancer Cell* **1,** 393–401 (2002).

131. Lin, P. P., Wang, Y. & Lozano, G. Mesenchymal Stem Cells and the Origin of Ewing's Sarcoma. *Sarcoma* **2011,** (2011).

132. Tirode, F. *et al.* Mesenchymal stem cell features of Ewing tumors. *Cancer Cell* **11,** 421–9 (2007).

133. Castillero-Trejo, Y., Eliazer, S., Xiang, L., Richardson, J. A. & Ilaria, R. L. Expression of the *EWS/FLI-1* Oncogene in Murine Primary Bone-Derived Cells Results in EWS/FLI-1–Dependent, Ewing Sarcoma–Like Tumors. *Cancer Res.* **65,** 8698–8705 (2005).

134. Riggi, N. *et al.* Development of Ewing's Sarcoma from Primary Bone Marrow–Derived Mesenchymal Progenitor Cells. *Cancer Res.* **65,** 11459–11468 (2005).

135. Miyagawa, Y. *et al.* Inducible expression of chimeric EWS/ETS proteins confers Ewing's family tumor-like phenotypes to human mesenchymal progenitor cells. *Mol. Cell. Biol.* **28,** 2125–37 (2008).

136. Riggi, N. *et al.* EWS-FLI-1 Expression Triggers a Ewing's Sarcoma Initiation Program in Primary Human Mesenchymal Stem Cells. *Cancer Res.* **68,** 2176–2185 (2008).

137. Cironi, L. *et al.* IGF1 Is a Common Target Gene of Ewing's Sarcoma Fusion Proteins in Mesenchymal Progenitor Cells. *PLoS One* **3,** e2634 (2008).

138. Weinstein, I. B. & Joe, A. Oncogene Addiction Multistage Carcinogenesis and Oncogene Addiction. *Cancer Res* **68,** 3077–80 (2008).

139. Weinstein, I. B. Disorders in cell circuitry during multistage carcinogenesis: the role of homeostasis. *Carcinogenesis* **21,** 857–64 (2000).

140. Weinstein, I. B. CANCER: Enhanced: Addiction to Oncogenes--the Achilles Heal

of Cancer. *Science (80-.).* **297,** 63–64 (2002).

141. Weinstein, I. B. & Joe, A. K. Mechanisms of Disease: oncogene addiction—a rationale for molecular targeting in cancer therapy. *Nat. Clin. Pract. Oncol.* **3,** 448–457 (2006).

142. Yan, C. & Higgins, P. J. Drugging the undruggable: Transcription therapy for cancer. *Biochim. Biophys. Acta - Rev. Cancer* **1835,** 76–85 (2013).

143. Koumenis, C. & Giaccia, A. Transformed cells require continuous activity of RNA polymerase II to resist oncogene-induced apoptosis. *Mol. Cell. Biol.* **17,** 7306–16 (1997).

144. Stellrecht, C. M. & Chen, L. S. Transcription inhibition as a therapeutic target for cancer. *Cancers (Basel).* **3,** 4170–90 (2011).

145. Villicaña, C., Cruz, G. & Zurita, M. The basal transcription machinery as a target for cancer therapy. *Cancer Cell Int.* **14,** 18 (2014).

146. Waddington, C. H. The Epigenotype. *Int. J. Epidemiol.* **41,** 10–13 (2012).

147. Sharma, S., Kelly, T. K. & Jones, P. A. Epigenetics in cancer. *Carcinogenesis* **31,** 27–36 (2010).

148. Goldberg, A. D., Allis, C. D. & Bernstein, E. Epigenetics: A Landscape Takes Shape. *Cell* **128,** 635–638 (2007).

149. Annunziato, A. DNA Packaging: Nucleosomes and Chromatin | Learn Science at Scitable. *Nature Education* (2008). doi:1(1):26

150. Luger, K., Mäder, A. W., Richmond, R. K., Sargent, D. F. & Richmond, T. J. Crystal structure of the nucleosome core particle at 2.8 Å resolution. *Nature* **389,** 251–260 (1997).

151. Dawson, M. A. & Kouzarides, T. Cancer epigenetics: from mechanism to therapy. *Cell* **150,** 12–27 (2012).

152. Robertson, K. D. DNA methylation and human disease. *Nat. Rev. Genet.* **6,** 597–610 (2005).

153. Wu, X. & Zhang, Y. TET-mediated active DNA demethylation: mechanism, function and beyond. *Nat. Rev. Genet.* **18,** 517–534 (2017).

154. Jin, B., Li, Y. & Robertson, K. D. DNA Methylation: Superior or Subordinate in the Epigenetic Hierarchy? *Genes Cancer* **2,** 607–617 (2011).

155. Portela, A. & Esteller, M. Epigenetic modifications and human disease. *Nat. Biotechnol.* **28,** 1057–1068 (2010).

156. Ley, T. J. *et al. DNMT3A* Mutations in Acute Myeloid Leukemia. *N. Engl. J. Med.* **363,** 2424–2433 (2010).

157. Allfrey, V. G., Faulkner, R. & Mirsky, A. E. Acetylation and methylation of histones and their possible role in the regulation of RNA synthesis. *Proc. Natl. Acad. Sci. U. S. A.* **51,** 786–94 (1964).

158. Lawrence, M., Daujat, S. & Schneider, R. Lateral Thinking: How Histone Modifications Regulate Gene Expression. (2016). doi:10.1016/j.tig.2015.10.007

159. Kouzarides, T. Chromatin Modifications and Their Function. *Cell* **128,** 693–705 (2007).

160. Shogren-Knaak, M. *et al.* Histone H4-K16 Acetylation Controls Chromatin Structure and Protein Interactions. *Science (80-.).* **311,** 844–847 (2006).

161. Lu, X. *et al.* The effect of H3K79 dimethylation and H4K20 trimethylation on nucleosome and chromatin structure. *Nat. Struct. Mol. Biol.* **15,** 1122–1124 (2008).

162. Strahl, B. D. & Allis, D. The language of covalent histone modifications. *Nat. www.nature.com* **403,** (2000).

163. Zhou, V. W., Goren, A. & Bernstein, B. E. Charting histone modifications and the functional organization of mammalian genomes. *Nat. Rev. Genet.* **12,** 7–18 (2011).

164. Baylin, S. B. & Jones, P. A. A decade of exploring the cancer epigenome — biological and translational implications. *Nat. Rev. Cancer* **11,** 726–734 (2011).

165. Becker, P. B. & Workman, J. L. Nucleosome Remodeling and Epigenetics. *Cold Spring Harb. Perspect. Biol.* **5,** a017905–a017905 (2013).

166. Flaus, A., Martin, D. M. A., Barton, G. J. & Owen-Hughes, T. Identification of multiple distinct SNF2 subfamilies with conserved structural motifs. *Nucleic Acids Res.* **34,** 2887–2905 (2006).

167. Clapier, C. R., Iwasa, J., Cairns, B. R. & Peterson, C. L. Mechanisms of action and regulation of ATP-dependent chromatin-remodelling complexes. *Nat. Rev. Mol. Cell Biol.* **18,** 407–422 (2017).

168. de la Serna, I. L., Ohkawa, Y. & Imbalzano, A. N. Chromatin remodelling in mammalian differentiation: lessons from ATP-dependent remodellers. *Nat. Rev. Genet.* **7,** 461–473 (2006).

169. Dürr, H., Flaus, A., Owen-Hughes, T. & Hopfner, K.-P. SNF2 family ATPases and DExx box helicases: differences and unifying concepts from high-resolution crystal structures. *Nucleic Acids Res.* **34,** 4160–7 (2006).

170. Ryan, D. P. & Owen-Hughes, T. SNF2-family proteins: chromatin remodellers for any occasion. *Curr. Opin. Chem. Biol.* **15,** 649–656 (2011).

171. Eisen, J. A., Sweder, K. S. & Hanawalt, P. C. Evolution of the SNF2 family of proteins: subfamilies with distinct sequences and functions. *Nucleic Acids Res.* **23,** 2715–23 (1995).

172. Clapier, C. R. & Cairns, B. R. The Biology of Chromatin Remodeling Complexes. *Annu. Rev. Biochem.* **78,** 273–304 (2009).

173. Medina, P. P., Sanchez-Cespedes, M. & Cespedes, M. S. Involvement of the chromatin-remodeling factor BRG1/SMARCA4 in human cancer. *Epigenetics* **3,** 64–8

174. Naidu, S. R., Love, I. M., Imbalzano, A. N., Grossman, S. R. & Androphy, E. J. The

SWI/SNF chromatin remodeling subunit BRG1 is a critical regulator of p53 necessary for proliferation of malignant cells. *Oncogene* **28,** 2492–2501 (2009).

175. Xue, Y. *et al.* NURD, a Novel Complex with Both ATP-Dependent Chromatin-Remodeling and Histone Deacetylase Activities. *Mol. Cell* **2,** 851–861 (1998).

176. Torchy, M. P., Hamiche, A. & Klaholz, B. P. Structure and function insights into the NuRD chromatin remodeling complex. *Cell. Mol. Life Sci.* **72,** 2491–2507 (2015).

177. Allen, H. F., Wade, P. A. & Kutateladze, T. G. The NuRD architecture. *Cell. Mol. Life Sci.* **70,** 3513–3524 (2013).

178. Basta, J. & Rauchman, M. The Nucleosome Remodeling and Deacetylase (NuRD) Complex in Development and Disease. doi:10.1016/j.trsl.2014.05.003

179. Le Guezennec, X. *et al.* MBD2/NuRD and MBD3/NuRD, Two Distinct Complexes with Different Biochemical and Functional Properties. *Mol. Cell. Biol.* **26,** 843–851 (2006).

180. Hendrich, B. & Bird, A. Identification and characterization of a family of mammalian methyl-CpG binding proteins. *Mol. Cell. Biol.* **18,** 6538–47 (1998).

181. Yildirim, O. *et al.* Mbd3/NURD complex regulates expression of 5-hydroxymethylcytosine marked genes in embryonic stem cells. *Cell* **147,** 1498–510 (2011).

182. Hendrich, B., Guy, J., Ramsahoye, B., Wilson, V. A. & Bird, A. Closely related proteins MBD2 and MBD3 play distinctive but interacting roles in mouse development. *Genes Dev.* **15,** 710–723 (2001).

183. Millard, C. J. *et al.* The structure of the core NuRD repression complex provides insights into its interaction with chromatin. *Elife* **5,** e13941 (2016).

184. Alqarni, S. S. M. *et al.* Insight into the architecture of the NuRD complex: structure of the RbAp48-MTA1 subcomplex. *J. Biol. Chem.* **289,** 21844–55 (2014).

185. Millard, C. J., Fairall, L. & Schwabe, J. W. R. Towards an understanding of the structure and function of MTA1. *Cancer Metastasis Rev.* **33,** 857–867 (2014).

186. Torrado, M. *et al.* Refinement of the subunit interaction network within the nucleosome remodelling and deacetylase (NuRD) complex. *FEBS J.* **284,** 4216–4232 (2017).

187. Hoffmeister, H. *et al.* CHD3 and CHD4 form distinct NuRD complexes with different yet overlapping functionality. *Nucleic Acids Res.* **45,** 10534–10554 (2017).

188. Ge, Q., Nilasena, D. S., O'Brien, C. A., Frank, M. B. & Targoff, I. N. Molecular analysis of a major antigenic region of the 240-kD protein of Mi-2 autoantigen. *J. Clin. Invest.* **96,** 1730–1737 (1995).

189. Flaus, A., Martin, D. M. A., Barton, G. J. & Owen-Hughes, T. Identification of multiple distinct SNF2 subfamilies with conserved structural motifs. *Nucleic Acids Res.* **34,** 2887–2905 (2006).

190. Hall, J. A. & Georgel, P. T. CHD proteins: a diverse family with strong ties.

Biochem. Cell Biol. **85,** 463–476 (2007).

191. Bouazoune, K. *et al.* The dMi-2 chromodomains are DNA binding modules important for ATP-dependent nucleosome mobilization. *EMBO J.* **21,** 2430–40 (2002).

192. Tencer, A. H. *et al.* Covalent Modifications of Histone H3K9 Promote Binding of CHD3. *Cell Rep.* **21,** 455–466 (2017).

193. Mansfield, R. E. *et al.* Plant Homeodomain (PHD) Fingers of CHD4 Are Histone H3-binding Modules with Preference for Unmodified H3K4 and Methylated H3K9. *J. Biol. Chem.* **286,** 11779–11791 (2011).

194. Musselman, C. A. *et al.* Binding of the CHD4 PHD2 finger to histone H3 is modulated by covalent modifications. *Biochem. J.* **423,** 179–187 (2009).

195. Watson, A. A. *et al.* The PHD and Chromo Domains Regulate the ATPase Activity of the Human Chromatin Remodeler CHD4. *J. Mol. Biol.* **422,** 3–17 (2012).

196. Morra, R., Lee, B. M., Shaw, H., Tuma, R. & Mancini, E. J. Concerted action of the PHD, chromo and motor domains regulates the human chromatin remodelling ATPase CHD4. *FEBS Lett.* **586,** 2513–2521 (2012).

197. Nitarska, J. *et al.* A Functional Switch of NuRD Chromatin Remodeling Complex Subunits Regulates Mouse Cortical Development. *Cell Rep.* **17,** 1683–1698 (2016).

198. Qi, W. *et al.* Acetyltransferase p300 collaborates with chromodomain helicase DNA-binding protein 4 (CHD4) to facilitate DNA double-strand break repair. *Mutagenesis* **31,** 193–203 (2016).

199. Smeenk, G. *et al.* The NuRD chromatin–remodeling complex regulates signaling and repair of DNA damage. *J. Cell Biol.* **190,** 741–749 (2010).

200. Polo, S. E., Kaidi, A., Baskcomb, L., Galanty, Y. & Jackson, S. P. Regulation of DNA-damage responses and cell-cycle progression by the chromatin remodelling factor CHD4. *EMBO J.* **29,** 3130–3139 (2010).

201. Lai, A. Y. & Wade, P. A. Cancer biology and NuRD: a multifaceted chromatin remodelling complex. *Nat. Rev. Cancer* **11,** 588–596 (2011).

202. Stanley, F. K. T., Moore, S. & Goodarzi, A. A. CHD chromatin remodelling enzymes and the DNA damage response. *Mutat. Res. Mol. Mech. Mutagen.* **750,** 31–44 (2013).

203. Helbling Chadwick, L., Chadwick, B. P., Jaye, D. L. & Wade, P. A. The Mi-2/NuRD complex associates with pericentromeric heterochromatin during S phase in rapidly proliferating lymphoid cells. *Chromosoma* **118,** 445–57 (2009).

204. Sims, J. K. & Wade, P. A. Mi-2/NuRD complex function is required for normal S phase progression and assembly of pericentric heterochromatin. *Mol. Biol. Cell* **22,** 3094–102 (2011).

205. Magdinier, F. & Wolffe, A. P. Selective association of the methyl-CpG binding protein MBD2 with the silent p14/p16 locus in human neoplasia. *Proc. Natl. Acad. Sci. U. S. A.* **98,** 4990–5 (2001).

206. Shimbo, T. *et al.* MBD3 Localizes at Promoters, Gene Bodies and Enhancers of Active Genes. *PLoS Genet.* **9,** e1004028 (2013).

207. Miccio, A. *et al.* NuRD mediates activating and repressive functions of GATA-1 and FOG-1 during blood development. *EMBO J.* **29,** 442–456 (2010).

208. Hosokawa, H. *et al.* Functionally distinct Gata3/Chd4 complexes coordinately establish T helper 2 (Th2) cell identity. *Proc. Natl. Acad. Sci. U. S. A.* **110,** 4691–6 (2013).

209. Reynolds, N. *et al.* NuRD-mediated deacetylation of H3K27 facilitates recruitment of Polycomb Repressive Complex 2 to direct gene repression. *EMBO J.* **31,** 593–605 (2012).

210. Gómez-del Arco, P. *et al.* The Chromatin Remodeling Complex Chd4/NuRD Controls Striated Muscle Identity and Metabolic Homeostasis. *Cell Metab.* **23,** 881–892 (2016).

211. Wilczewski, C. M. *et al.* CHD4 and the NuRD complex directly control cardiac sarcomere formation. *Proc. Natl. Acad. Sci.* 201722219 (2018). doi:10.1073/pnas.1722219115

212. Sparmann, A. *et al.* The chromodomain helicase Chd4 is required for Polycomb-mediated inhibition of astroglial differentiation. *EMBO J.* **32,** 1598–1612 (2013).

213. Muralidharan, B. *et al.* LHX2 Interacts with the NuRD Complex and Regulates Cortical Neuron Subtype Determinants Fezf2 and Sox11. *J. Neurosci.* **37,** 194–203 (2017).

214. Reynolds, N. *et al.* NuRD Suppresses Pluripotency Gene Expression to Promote Transcriptional Heterogeneity and Lineage Commitment. *Cell Stem Cell* **10,** 583–594 (2012).

215. Zhao, H. *et al.* The chromatin remodeler Chd4 maintains embryonic stem cell identity by controlling pluripotency- and differentiation-associated genes. *J. Biol. Chem.* **292,** 8507–8519 (2017).

216. Le Gallo, M. *et al.* Exome sequencing of serous endometrial tumors identifies recurrent somatic mutations in chromatin-remodeling and ubiquitin ligase complex genes. *Nat. Genet.* **44,** 1310–1315 (2012).

217. Malisetty, V. L., Penugurti, V., Panta, P., Chitta, S. K. & Manavathi, B. MTA1 expression in human cancers – Clinical and pharmacological significance. *Biomed. Pharmacother.* **95,** 956–964 (2017).

218. Dong, H. *et al.* The Metastasis-Associated Gene MTA3, a Component of the Mi-2/NuRD Transcriptional Repression Complex, Predicts Prognosis of Gastroesophageal Junction Adenocarcinoma. *PLoS One* **8,** e62986 (2013).

219. Zhang, X. -y. *et al.* Metastasis-associated protein 1 (MTA1) is an essential downstream effector of the c-MYC oncoprotein. *Proc. Natl. Acad. Sci.* **102,** 13968–13973 (2005).

220. Ginder, G. D. & Williams, D. C. Readers of DNA methylation, the MBD family as

potential therapeutic targets. *Pharmacol. Ther.* **184,** 98–111 (2018).

221. Ma, C. *et al.* SALL1 functions as a tumor suppressor in breast cancer by regulating cancer cell senescence and metastasis through the NuRD complex. *Mol. Cancer* **17,** 78 (2018).

222. Dhanak, D. & Jackson, P. Development and classes of epigenetic drugs for cancer. *Biochem. Biophys. Res. Commun.* **455,** 58–69 (2014).

223. Johnston, S. J. & Carroll, J. S. Transcription factors and chromatin proteins as therapeutic targets in cancer. *Biochim. Biophys. Acta - Rev. Cancer* **1855,** 183–192 (2015).

224. Helin, K. & Dhanak, D. Chromatin proteins and modifications as drug targets. *Nature* **502,** 480–488 (2013).

225. Okada, Y. *et al.* hDOT1L Links Histone Methylation to Leukemogenesis. *Cell* **121,** 167–178 (2005).

226. Filippakopoulos, P. *et al.* Selective inhibition of BET bromodomains. *Nature* **468,** 1067–1073 (2010).

227. Delmore, J. E. *et al.* BET Bromodomain Inhibition as a Therapeutic Strategy to Target c-Myc. *Cell* **146,** 904–917 (2011).

228. Puissant, A. *et al.* Targeting MYCN in Neuroblastoma by BET Bromodomain Inhibition. *Cancer Discov.* **3,** 308–323 (2013).

229. DeWoskin, V. A. & Million, R. P. The epigenetics pipeline. *Nat. Rev. Drug Discov.* **12,** 661–662 (2013).

230. Lee, T. I. & Young, R. A. Transcriptional Regulation and Its Misregulation in Disease. *Cell* **152,** 1237–1251 (2013).

231. Nambiar, M., Kari, V. & Raghavan, S. C. Chromosomal translocations in cancer. *Biochim. Biophys. Acta - Rev. Cancer* **1786,** 139–152 (2008).

232. Sur, I. & Taipale, J. The role of enhancers in cancer. *Nat. Rev. Cancer* **16,** 483–493 (2016).

233. Cao, L. *et al.* Genome-wide identification of PAX3-FKHR binding sites in rhabdomyosarcoma reveals candidate target genes important for development and cancer. *Cancer Res.* **70,** 6497–508 (2010).

234. Dukler, N., Gulko, B., Huang, Y.-F. & Siepel, A. Is a super-enhancer greater than the sum of its parts? *Nat. Genet.* **49,** 2–3 (2017).

235. Sengupta, S. & George, R. E. Super-Enhancer-Driven Transcriptional Dependencies in Cancer. *Trends in Cancer* **3,** 269–281 (2017).

236. Whyte, W. A. *et al.* Master transcription factors and mediator establish super-enhancers at key cell identity genes. *Cell* **153,** 307–19 (2013).

237. de Dieuleveult, M. *et al.* Genome-wide nucleosome specificity and function of chromatin remodellers in ES cells. *Nature* **530,** 113–116 (2016).

238. Miller, A. *et al.* Sall4 controls differentiation of pluripotent cells independently of the Nucleosome Remodelling and Deacetylation (NuRD) complex. *Development* **143,** 3074–84 (2016).

239. Shimbo, T. *et al.* MBD3 Localizes at Promoters, Gene Bodies and Enhancers of Active Genes. *PLoS Genet.* **9,** e1004028 (2013).

240. Hnisz, D. *et al.* Super-enhancers in the control of cell identity and disease. *Cell* **155,** 934–47 (2013).

241. Bornelöv, S. *et al.* The Nucleosome Remodeling and Deacetylation Complex Modulates Chromatin Structure at Sites of Active Transcription to Fine-Tune Gene Expression. *Mol. Cell* **71,** 56–72.e4 (2018).

242. Riggi, N. *et al.* EWS-FLI1 utilizes divergent chromatin remodeling mechanisms to directly activate or repress enhancer elements in Ewing sarcoma. *Cancer Cell* **26,** 668–681 (2014).

243. Boulay, G. *et al.* Cancer-Specific Retargeting of BAF Complexes by a Prion-like Domain. *Cell* **171,** 163–178.e19 (2017).

244. Gollavilli, P. N. *et al.* EWS/ETS-driven Ewing Sarcoma requires BET bromodomain proteins. *Cancer Res.* canres.0484.2018 (2018). doi:10.1158/0008-5472.CAN-18-0484

245. Tomazou, E. M. *et al.* Epigenome Mapping Reveals Distinct Modes of Gene Regulation and Widespread Enhancer Reprogramming by the Oncogenic Fusion Protein EWS-FLI1 Article Epigenome Mapping Reveals Distinct Modes of Gene Regulation and Widespread Enhancer Reprogramming by the Oncogenic Fusion Protein EWS-FLI1. *Cell Rep.* **10,** (2015).

246. Chipumuro, E. *et al.* CDK7 inhibition suppresses super-enhancer-linked oncogenic transcription in MYCN-driven cancer. *Cell* **159,** 1126–1139 (2014).

247. Christensen, C. L. *et al.* Targeting transcriptional addictions in small cell lung cancer with a covalent CDK7 inhibitor. *Cancer Cell* **26,** 909–922 (2014).

248. Hnisz, D. *et al.* Convergence of Developmental and Oncogenic Signaling Pathways at Transcriptional Super-Enhancers. *Mol. Cell* **58,** 362–370 (2015).

249. Yokoyama, Y. *et al.* BET Inhibitors Suppress ALDH Activity by Targeting *ALDH1A1* Super-Enhancer in Ovarian Cancer. *Cancer Res.* **76,** 6320–6330 (2016).

250. Bhagwat, A. S. *et al.* BET Bromodomain Inhibition Releases the Mediator Complex from Select cis -Regulatory Elements. *Cell Rep.* **15,** 519–530 (2016).

251. Chapuy, B. *et al.* Discovery and characterization of super-enhancer-associated dependencies in diffuse large B cell lymphoma. *Cancer Cell* **24,** 777–90 (2013).

252. Lovén, J. *et al.* Selective inhibition of tumor oncogenes by disruption of super-enhancers. *Cell* **153,** 320–34 (2013).

253. Jacques, C. *et al.* Targeting the epigenetic readers in Ewing sarcoma inhibits the oncogenic transcription factor EWS/Fli1. *Oncotarget* **7,** 24125–40 (2016).

254. Hensel, T. *et al.* Targeting the EWS-ETS transcriptional program by BET bromodomain inhibition in Ewing sarcoma. *Oncotarget* **7,** 1451–63 (2016).

255. Loganathan, S. N. *et al.* BET bromodomain inhibitors suppress EWS-FLI1-dependent transcription and the IGF1 autocrine mechanism in Ewing sarcoma. *Oncotarget* **7,** 43504–43517 (2016).

256. Chudnovsky, Y. *et al.* ZFHX4 Interacts with the NuRD Core Member CHD4 and Regulates the Glioblastoma Tumor-Initiating Cell State. *Cell Rep.* **6,** 313–324 (2014).

257. Wai, D. C. C. *et al.* The BRD3 ET domain recognizes a short peptide motif through a mechanism that is conserved across chromatin remodelers and transcriptional regulators. *J. Biol. Chem.* **293,** 7160–7175 (2018).

258. Winter, G. E. *et al.* BET Bromodomain Proteins Function as Master Transcription Elongation Factors Independent of CDK9 Recruitment. *Mol. Cell* **67,** 5–18.e19 (2017).

259. de Dieuleveult, M. *et al.* Genome-wide nucleosome specificity and function of chromatin remodellers in ES cells. *Nature* **530,** 113–116 (2016).

260. Xia, L. *et al.* CHD4 Has Oncogenic Functions in Initiating and Maintaining Epigenetic Suppression of Multiple Tumor Suppressor Genes. *Cancer Cell* **31,** 653–668.e7 (2017).

261. Cai, Y. *et al.* The NuRD complex cooperates with DNMTs to maintain silencing of key colorectal tumor suppressor genes. *Oncogene* **33,** 2157–2168 (2014).

262. O'Shaughnessy-Kirwan, A., Signolet, J., Costello, I., Gharbi, S. & Hendrich, B. Constraint of gene expression by the chromatin remodelling protein CHD4 facilitates lineage specification. *Development* **142,** 2586–2597 (2015).

263. Zhang, J. *et al.* Harnessing of the nucleosome-remodeling-deacetylase complex controls lymphocyte development and prevents leukemogenesis. *Nat. Immunol.* **13,** 86–94 (2012).

264. Wang, Y. *et al.* LSD1 is a subunit of the NuRD complex and targets the metastasis programs in breast cancer. *Cell* **138,** 660–72 (2009).

265. Low, J. K. K. *et al.* CHD4 Is a Peripheral Component of the Nucleosome Remodeling and Deacetylase Complex. *J. Biol. Chem.* **291,** 15853–66 (2016).

266. Hedrick, E., Crose, L., Linardic, C. M. & Safe, S. Histone Deacetylase Inhibitors Inhibit Rhabdomyosarcoma by Reactive Oxygen Species-Dependent Targeting of Specificity Protein Transcription Factors. *Mol. Cancer Ther.* **14,** 2143–2153 (2015).

267. Enßle, J. C. *et al.* Co-targeting of BET proteins and HDACs as a novel approach to trigger apoptosis in rhabdomyosarcoma cells. *Cancer Lett.* **428,** 160–172 (2018).

268. Haydn, T., Metzger, E., Schuele, R. & Fulda, S. Concomitant epigenetic targeting of LSD1 and HDAC synergistically induces mitochondrial apoptosis in rhabdomyosarcoma cells. *Cell Death Dis.* **8,** e2879 (2017).

269. Zhang, W. *et al.* The Nucleosome Remodeling and Deacetylase Complex NuRD Is Built from Preformed Catalytically Active Sub-modules. *J. Mol. Biol.* **428,** 2931–

2942 (2016).

270. Topark-Ngarm, A. *et al.* CTIP2 associates with the NuRD complex on the promoter of p57KIP2, a newly identified CTIP2 target gene. *J. Biol. Chem.* **281,** 32272–83 (2006).

271. Ramírez, J., Dege, C., Kutateladze, T. G. & Hagman, J. MBD2 and multiple domains of CHD4 are required for transcriptional repression by Mi-2/NuRD complexes. *Mol. Cell. Biol.* **32,** 5078–88 (2012).

272. Mohd-Sarip, A. *et al.* DOC1-Dependent Recruitment of NURD Reveals Antagonism with SWI/SNF during Epithelial-Mesenchymal Transition in Oral Cancer Cells. *Cell Rep.* **20,** 61–75 (2017).

273. Zabidi, M. A. *et al.* Enhancer–core-promoter specificity separates developmental and housekeeping gene regulation. *Nature* **518,** 556–559 (2015).

274. Zabidi, M. A. & Stark, A. Regulatory Enhancer-Core-Promoter Communication via Transcription Factors and Cofactors. *Trends Genet.* **32,** 801–814 (2016).

275. Xie, W. *et al.* The chromatin remodeling complex NuRD establishes the poised state of rRNA genes characterized by bivalent histone modifications and altered nucleosome positions. *Proc. Natl. Acad. Sci. U. S. A.* **109,** 8161–6 (2012).

276. Souza, B. K. *et al.* Targeting Histone Deacetylase Activity to Arrest Cell Growth and Promote Neural Differentiation in Ewing Sarcoma. *Mol. Neurobiol.* (2018). doi:10.1007/s12035-018-0874-6

277. Di Pompo, G. *et al.* Novel Histone Deacetylase Inhibitors Induce Growth Arrest, Apoptosis, and Differentiation in Sarcoma Cancer Stem Cells. *J. Med. Chem.* **58,** 4073–4079 (2015).

278. Pattenden, S. G. *et al.* High-throughput small molecule screen identifies inhibitors of aberrant chromatin accessibility. *Proc. Natl. Acad. Sci.* **113,** 3018–3023 (2016).

279. Zorzi, A. P. *et al.* A phase I study of histone deacetylase inhibitor, pracinostat (SB939), in pediatric patients with refractory solid tumors: IND203 a trial of the NCIC IND program/C17 pediatric phase I consortium. *Pediatr. Blood Cancer* **60,** 1868–1874 (2013).

280. Li, Y. *et al.* Impairment of p53 acetylation by EWS-Fli1 chimeric protein in Ewing Family Tumors. *Cancer Lett.* **320,** 14–22 (2012).

281. Sankar, S. *et al.* Reversible LSD1 inhibition interferes with global EWS/ETS transcriptional activity and impedes Ewing sarcoma tumor growth. *Clin. Cancer Res.* **20,** 4584–97 (2014).

282. D'Alesio, C. *et al.* RNAi screens identify CHD4 as an essential gene in breast cancer growth. *Oncotarget* **7,** 80901–80915 (2016).

283. Heshmati, Y. *et al.* Identification of CHD4 As a Potential Therapeutic Target of Acute Myeloid Leukemia. *Blood* **128,** (2016).

284. Kwan, A. H. Y. *et al.* Engineering a Protein Scaffold from a PHD Finger. *Structure* **11,** 803–813 (2003).

285. Musselman, C. A. *et al.* Bivalent recognition of nucleosomes by the tandem PHD fingers of the CHD4 ATPase is required for CHD4-mediated repression. *Proc. Natl. Acad. Sci. U. S. A.* **109,** 787–92 (2012).

286. Musselman, C. A. & Kutateladze, T. G. PHD Fingers: Epigenetic Effectors and Potential Drug Targets. *Mol. Interv.* **9,** 314–323 (2009).

287. Baker, L. A., Allis, C. D. & Wang, G. G. PHD fingers in human diseases: disorders arising from misinterpreting epigenetic marks. *Mutat. Res.* **647,** 3–12 (2008).

288. Musco, G. & Peterson, P. PHD finger of autoimmune regulator: an epigenetic link between the histone modifications and tissue-specific antigen expression in thymus. *Epigenetics* **3,** 310–4 (2008).

289. Santiago, C., Nguyen, K. & Schapira, M. Druggability of methyl-lysine binding sites. *J. Comput. Aided. Mol. Des.* **25,** 1171–8 (2011).

290. Miller, T. C. R. *et al.* Competitive binding of a benzimidazole to the histone-binding pocket of the Pygo PHD finger. *ACS Chem. Biol.* **9,** 2864–74 (2014).

291. Wagner, E. K., Nath, N., Flemming, R., Feltenberger, J. B. & Denu, J. M. Identification and Characterization of Small Molecule Inhibitors of a Plant Homeodomain Finger. *Biochemistry* **51,** 8293–8306 (2012).

292. Wang, T. *et al.* Identification and characterization of essential genes in the human genome. *Science* **350,** 1096–101 (2015).

293. Erez-Salvia, M. P. & Esteller, M. Bromodomain inhibitors and cancer therapy: From structures to applications. *Epigenetics* **12,** (2017).

294. Berthon, C. *et al.* Bromodomain inhibitor OTX015 in patients with acute leukaemia: a dose-escalation, phase 1 study. *Lancet Haematol.* **3,** e186–e195 (2016).

295. Amorim, S. *et al.* Bromodomain inhibitor OTX015 in patients with lymphoma or multiple myeloma: a dose-escalation, open-label, pharmacokinetic, phase 1 study. *Lancet Haematol.* **3,** e196–e204 (2016).

296. Clark, J. *et al.* Identification of novel genes, SYT and SSX, involved in the t(X;18)(p11.2;q11.2) translocation found in human synovial sarcoma. *Nat. Genet.* **7,** 502–508 (1994).

297. McBride, M. J. *et al.* The SS18-SSX Fusion Oncoprotein Hijacks BAF Complex Targeting and Function to Drive Synovial Sarcoma. *Cancer Cell* **33,** 1128–1141.e7 (2018).

298. Kadoch, C. & Crabtree, G. R. Reversible Disruption of mSWI/SNF (BAF) Complexes by the SS18-SSX Oncogenic Fusion in Synovial Sarcoma. *Cell* **153,** 71–85 (2013).

299. McDonald, E. R. *et al.* Project DRIVE: A Compendium of Cancer Dependencies and Synthetic Lethal Relationships Uncovered by Large-Scale, Deep RNAi Screening. *Cell* **170,** 577–592.e10 (2017).

300. Chudnovsky, Y. *et al.* ZFHX4 Interacts with the NuRD Core Member CHD4 and Regulates the Glioblastoma Tumor-Initiating Cell State. *Cell Rep.* **6,** 313–324 (2014).

301. Nio, K. *et al.* Defeating EpCAM+ liver cancer stem cells by targeting chromatin remodeling enzyme CHD4 in human hepatocellular carcinoma. *J. Hepatol.* **63,** 1164–1172 (2015).

302. Xu, N., Liu, F., Zhou, J. & Bai, C. CHD4 is associated with poor prognosis of non-small cell lung cancer patients through promoting tumor cell proliferation. in *11.1 Lung Cancer* **48,** PA2862 (European Respiratory Society, 2016).

303. Yamada, M. *et al.* Association of the chromodomain helicase DNA-binding protein 4 (*CHD4*) missense variation p.D140E with cancer: Potential interaction with smoking. *Genes, Chromosom. Cancer* **54,** 122–128 (2015).

304. Zhang, L. *et al.* EZH2-, CHD4-, and IDH-linked epigenetic perturbation and its association with survival in glioma patients. *J. Mol. Cell Biol.* **9,** 477–488 (2017).

305. Sperlazza, J. *et al.* Depletion of the chromatin remodeler CHD4 sensitizes AML blasts to genotoxic agents and reduces tumor formation. *Blood* **126,** 1462–1472 (2015).

Acknowledgments

First, I would like to thank my thesis supervisor Beat Schäfer for offering me such an exciting project, for trusting my judgement, for always having the door open for discussion, and for allowing me to explore my ideas.

I am thankful to all my Committee Members. To Raffaella Santoro, for welcoming me to her laboratory, for the nice discussions, and for being interested in my development. To Amedeo Caflisch for also welcoming me to his laboratory, for the many meetings, and for pushing the CHD4 inhibitor development. To César Fernández for always being available, for always greeting me with a smile, for inviting me to work at Novartis, for helping me throughout the many PhD applications, and for being such a great mentor and person.

I am also very thankful to Javed Khan, who did not hesitate in granting me the opportunity to change country and continent and pursue my research at the National Cancer Institute – NIH. Going to America was an important turning point in my PhD. I grew and learnt a lot during those 6 months in Bethesda. I was incredibly welcomed in the laboratory of Javed Khan and happy to have met and worked with such great scientists. A special thanks goes to Berkley Gryder, my mentor, companion in the epigenetic dissection mission, and above all my friend. I am also very happy to have met the brilliant mind and sweet sole of Marielle Yohe. I could not finish this section without thanking Young Song, the man who can sequence DNA out of water, for his kindness, Nitya Shivaprasad for her companionship and Sofia Girald for not letting me spend Christmas alone.

A special thank you goes to the groups of Raffaella and Amedeo for making me feel that I was also a part of the lab. Particularly, I want to thank Dimitrios Spiliotopoulos for the nice discussions and for his efforts in the pursue of a CHD4 inhibitor.

I am incredibly grateful for all the past and current members of the sarcoma laboratory. It is really a blessing to work with such nice and warm people.
To Marco Wachtel, you know the lab would not be the same without you, thank you for the discussions, for knowing all the projects, for never getting tired of endless and repetitive questions, for all the papers and grants' corrections, and for all the nice jokes.
A very special thank you goes to Maria Böhm. I could thank you for giving me such a nice project, but I rather thank you for your friendship, your support, and your hugs.

Very special thank yous go to my amazing master students, Yeonjoo Chung and Blaz Pavlovic. I could never have wished for more dedicated, friendly, hard-working, and sweet students. I was very blessed. A lot of the work shown here is thanks to Blaz Pavlovic who worked tirelessly for uncountable hours with me for 1 year and half. Thank you for all the nice moments inside and outside the lab.
To Johannes, with whom I shared these 4 years and half (minus 2 weeks), I want to thank you for providing me some of your calmness and comfort in times of need; to Eva Brack I want thank you for being our on-call energetic doctor ready to answer our every need; to Dominik Laubscher I want to thank you for your easiness and for providing such a peaceful atmosphere; to the forever cold Gabriele Manzella I want to thank you for your insights; to Chiara Giorgi I want to thank you for keeping everything and everybody in order; to Michaela Römmele I want to thank you for your fantastic chocolate drawer; to Quy Ngo, I want to thank you for your dedication; to Katharina Holste, my dearest schatz, I want to thank you for being this force of nature; to Nagjie Alijaj and her contagious laugh I want to thank you for being a sunshine in our lab; to Leonie Schreck I want to thank you for being my favorite büsi acquaintance; to Maria Gierisch I want to thank you for always being ready to help; to Vadim Saratov I want to thank you for your endless trivia; to Irina Banzola I want to thank you for making sure we have what we need; to Devmini Moonamale I want to welcome to our very special lab.
Also, I want to thank all the EICR, past and current, members for all the good moments, the helpful discussions, and the companionship.

I am very very very grateful for my friends.
A special thank you goes to Gloria Pedot. It was really amazing to share these last two years with you. Do not know what say except for thank you for taking care of me.
I want to thank all my friends in Zurich and Basel, who never made me regret leaving Portugal and embracing this opportunity. Specially, I want to thank my girls Filipa Fonseca and Marion Bailly. You make my life easier, nicer, and full of love. Also want to thank my papa André Cascais, I know you care for me too. To my friends in Portugal, I am thankful for making the distance so small. Specially want to thank my travel buddies Anna Cescon and Marta Santos, I would lose my mind without our yearly trips.

E por fim, quero agradecer à minha Família, não pelos 4 anos que passaram, mas por todos os 29. Aquilo que sou devo-o a vós. Mãe, Pai, Sofia, João, Tio, Avó e Avô, obrigado por verem em mim o que eu não vejo, por me terem criado e aturado, pelo amor, por acreditarem em mim, por me fazerem tão tão feliz. Tenho saudades vossas todos os dias da minha vida.

Chapter Five: Annexes

List of abbreviations

ALL – acute lymphoblastic leukemia

AML – acute myeloid leukemia

aRMS – alveolar rhabdomyosarcoma

BET – bromodomain and extra-terminal

bp – base pairs

BRD4 – bromodomain-containing protein 4

CHD4 – chromodomain-helicase-DNA binding protein 4

ChIP-seq – chromatin immunoprecipitation and sequencing

CML – chronic myeloid leukemia

CT – computed tomography

DDR – DNA-damage repair

DMEM – Dulbecco's modified Eagle's medium

DNA – deoxyribonucleic acid

DNase-HS – DNase hypersensitivity

DNMT – DNA-methyl transferase

DSS – 4,4-dimethyl-4-silapentane-1-sulfonic acid

EMT – epithelial-mesenchymal transition

eRMS – embryonal rhabdomyosarcoma

ES – Ewing sarcoma

ESC – embryonic stem cell

ET – extra-terminal

FACS – fluorescence-activated cell sorting

FISH – fluorescence in situ hybridization

FN-RMS – fusion-negative rhabdomyosarcoma

FP-RMS – fusion-positive rhabdomyosarcoma

GSEA – gene set enrichment analysis

H3 – histone 3

HDAC – histone deacetylase

HE – hematoxylin and eosin

HMT – histone methyltransferase

HSQC – heteronuclear single quantum coherence

IHC – immunohistochemistry

IPTG – isopropyl-thiogalactopyranoside

LOH – loss of heterozygosity

LOI – loss of imprinting
LSD1 – lysine-specific histone demethylase 1A
MBT – malignant brain tumor
MRI – magnetic resonance imaging
MSC – mesenchymal stem cell
MTF – master transcription factor
NPC – neural progenitor cell
NuRD – nucleosome remodeling and deacetylase complex
P3F – PAX3-FOXO1
PET – positron emission tomography
PFA – paraformaldehyde
PHD – plant homeodomain
PRC2 – Polycomb protein complex
P-TEFb – positive elongation factor complex b
PTM – post-translational modification
qPCR – Quantitative polymerase chain reaction
RMS – rhabdomyosarcoma
RNA – ribonucleic acid
RPMI – Roswell Park Memorial Institute
rRNA – ribosomal RNA
RT-PCR – real-time polymerase chain reaction
SE – super-enhancer
SEC – size exclusion chromatography
shRNA – short hairpin RNA
sgRNA – Guide RNA
TEV – tobacco etch virus
TF – transcription factor
TSS – transcription start site
WHO – world health organization

Curriculum Vitae

Personal information

Name: Joana Graça Marques
Address: Gladbachstrasse 94, 8044 Zurich, Switzerland
Mobile phone: +41 76 263 1258
Date of birth: 08-May-1989
E-mail: joanagracamarques@gmail.com
Nationality: Portuguese

Professional experience

April 2014 – Current date	PhD Student University Children's Hospital, Zurich, Switzerland Cancer Biology PhD program Life Science Zurich Graduate School Subject: CHD4 as a new target for super-enhancer disruption therapy in pediatric sarcoma Supervision: Prof. Beat Schäfer (PhD)
August 2015 – Jan 2016	Visiting PhD Student National Cancer Institute-NIH, Washington DC, USA Activities: ChIP-seq and DNase hypersensitivity assays Supervision: Javed Khan (MD)
February 2013 – Jan 2014	Structural Biophysics trainee Novartis Institutes for Biomedical Research, Basel, Switzerland Activities: NMR, SPR. and crystallography Supervision: César Fernández (PhD)

Teaching experience

January 2017 – Jan 2018	Master thesis Supervisor Master in Cancer Biology, University of Zurich, Switzerland
July 2016 – Jan 2017	Master thesis Supervisor

	Master in Medicinal and Industrial Pharmaceutical Sciences, ETH, Zurich, Switzerland
January 2010 – Dec 2011	Tutor Faculty of Pharmacy, University of Lisbon, Portugal Classes assisted: Analytical Chemistry, Instrumental Methods of Analysis and Pharmaceutical Chemistry

Education

2007 – 2013	Master Degree in Pharmaceutical Sciences Faculty of Pharmacy, University of Lisbon, Portugal Qualification: 17 out of 20 (Very Good)

Research Work

April 2012 – June 2012	Title: "Identification of new liver fibrosis biomarkers" School of Pharmacy, University College London Supervision: Rosemary Smyth (PhD) Qualification: 19 out of 20 (Excellent)
Sept 2011 – Feb 2012	Title: "Methodologies for primary culture hepatocytes biotransformation evaluation" CBT iMed.UL/Faculty of Pharmacy (FFUL) Supervision: Joana Miranda (PhD) Qualification: 18 out of 20 (Excellent)
Sept 2010 – Feb 2011	Title: "Evaluation of cadmium cytotoxicity in human breast cells" CBT iMed.UL/Faculty of Pharmacy (FFUL) Supervision: Prof. Nuno Oliveira (PhD) Qualification: 19 out of 20 (Excellent)
Sept 2009 – Oct 2010	Title: "Modulation of HNE cytotoxicity" Amadeu Dias Scholarship at CBT iMed.UL/FFUL Supervision: Prof. Nuno Oliveira (PhD)

Awards and Grants

Sept 2017 – June 2018	Forschungskredit Candoc 2017 Funding granted by the University of Zurich Salary funding for the 4th year of my PhD
2017	Travel grant Funding granted by the Cancer Biology PhD Program, University of Zurich
March 2016	Poster price 6th Cancer Biology PhD Student Retreat Held in Davos, Switzerland
2013	School Incentive Award for Postgraduate Training Granted by the Faculty of Pharmacy, University of Lisbon
2009 - 2010	Amadeu Dias Foundation Scholarship Granted by the Amadeus Dias Foundation Funding for a 1-year research project

Public Communications

November 2017	"ATP-dependent chromatin remodelling complexes as valuable new targets for pediatric cancer treatment" Talk at the Colloquium of Applied Cancer Research Held by the University Hospital Zurich in Zurich, Switzerland
October 2017	"Targeting chromatin accessibility as a potential new therapeutic strategy for pediatric sarcoma" Selected talk at the 7th FZK/CRC Retreat Held by the Kinderspital Zurich in Zurich, Switzerland
February 2017	"ATP-dependent chromatin remodelling complexes as valuable new targets for pediatric cancer treatment" Talk at Joint Cancer Meeting, Cancer Network Zurich Zurich, Switzerland

April 2016	"Epigenetics in rhabdomyosarcoma: the NuRD complex as a coregulator of PAX3-FOX1 mediated expression" Talk at the Colloquium of Applied Cancer Research Held by the University Hospital Zurich in Zurich, Switzerland
November 2015	"Chromatin remodelers: possible new therapeutic targets for fusion positive rhabdomyosarcoma" Selected Talk at the conference – "Advances in Pediatric Cancer Research: From Mechanisms and Models to Treatment and Survivorship" Held by AACR in Fort Lauderdale, Florida, USA

Competences and Interests

Technical	FACS, ChIP-seq, RNA-seq, NGS, DNase hypersensitivity
Computer	Basic R, Advanced in Microsoft Office (Word, Excel, Powerpoint), Intermediate in Adobe Creative Suite (Photoshop, Illustrator), GraphPad Prism
Languages	Portuguese: Native speaker English: Proficiency French: Independent speaker German: Currently learning (A2 level)
Interests	Theatre, drawing, oil painting, salsa dancing and travelling

List of Publications

Gryder BE, Yohe ME, Chou H, Zhang X, **Marques JG**, Wachtel M, Schäfer BW, Sen N, Song YK, Gualtieri A, Pomella S, Rota R, Cleveland A, Wen X, Sindir Si, Wei JS, Barr FG, Das S, Andresson T, Guha R, Lal-Nag M, Ferrer M, Shern JF, Zhao K, Thomas CJ, Khan J; PAX3-FOXO1 Establishes Myogenic Super Enhancers and Confers BET Bromodomain Vulnerability; *Cancer Discovery.* Aug;7(8):884-899 (2017).

Böhm M, Wachtel M, **Marques JG**, Laubscher D, Streiff N, Nanni P, Mamchaoui K, Santoro R, Schäfer BW; Helicase CHD4 is an epigenetic coregulator of PAX3-FOXO1 in alveolar rhabdomyosarcoma; *Journal of Clinical Investigation*. 126(11):4237-4249 (2016).

Gonçalves S, Fernandes AS, Oliveira NG, **Marques JG**, Costa J, Cabral MF, Miranda J, Cipriano M, Guerreiro PS, Castro M; Cytotoxic effects of cadmium in mammary epithelial cells: Protective role of the macrocycle [15]pyN5; *Food ChemToxicol.* 50(6):2180-7(2012).